PFUND ISSUE

Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 37, NUMBER 10 OCTOBER, 1947

The Use of Echelette Gratings in High Orders

R. W. WOOD
Johns Hopkins University, Baltimore, Maryland
(Received July 18, 1947)

AT the annual meeting of the Optical Society in October, 1946, a brief account was given of some preliminary experiments, made with improvised apparatus at East Hampton in my old barn laboratory, along lines suggested by Dr. Shane of the Lick Observatory. The idea was to employ a reflecting echelette grating of the type used for investigations in the infra-red region and concentrating in very high orders for visible light, crossed with a prism for the separation of the superposed orders of spectra. Such a combination, if it could be made of sufficient efficiency, might be employed in the construction of a stellar spectrograph in which a long spectrum of high dispersion and resolving power could be "folded up" so to speak into a compact group of superposed short spectra of high order, one above another. The advantage of such a disposition would be that all parts of a spectrum of high dispersion could be imaged in the immediate vicinity of the optic axis of the camera lens or mirror, and that a photographic plate, say two feet in length, which would be required for recording a spectrum of similar dispersion by the customary optical systems, could be replaced by one having a width of less than two inches.

In order to get a picture of the field of operations, which would serve also as a map on which the positions of faint or absent orders, as called for by theory could be recorded, I first made a plot on coordinate paper of the spectra of all orders between the first and thirtieth, arranged close together one above another. A plot of the spectra between the first and the eighth order will serve, however, for the demonstration of the position of, and relations between the superposed orders of spectra and the conditions which must be fulfilled in their employment in practice, as well as the minimum width of the plate or film which can be employed for the registration of the complete spectrum between any two given wave-lengths. It shows also that this minimum plate width is the same in the regions of very low and very high orders. Such a plot is reproduced in Fig. 1 for the region between the first and the eighth orders. We see at once that the length of the spectrum in any order is equal to the length of the first order spectrum multiplied by the order number. Yet, as has been said, the minimum width for recording a spectrum of any given length, when operating in the region of very high orders, is the same as when the lowest orders are employed. We find also that a line of any order in the corridor, say $\lambda = 4000$ in the seventh order, coincides with a point in the order below represented by the relation (4000

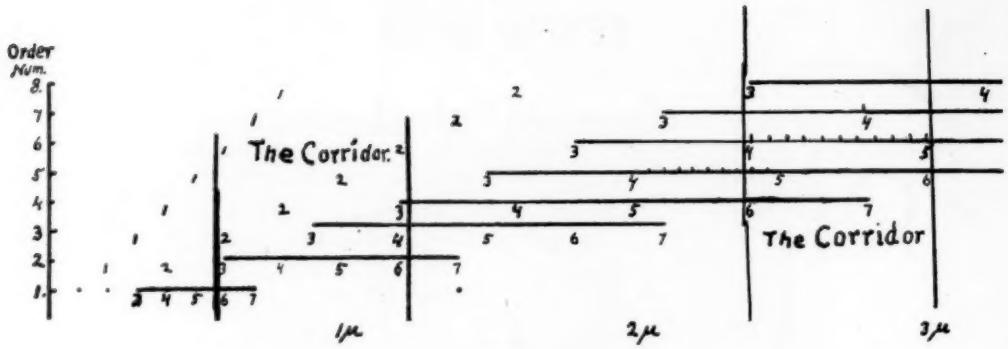

FIG. 1.

$\times 7)/6 = \lambda 4666$, $(4000 \times 6)/5 = 4800$, and $(4000 \times 6)/4 = 6000$. Employment of these relations will be made later on.

Referring now to Fig. 1, we select the points recording $\lambda = 6000$ in the fourth and fifth orders, and erect two vertical lines embracing these points. As shown, we find that all wave-lengths between 3000 and 7500 will be included in the region between these lines, which region I shall call the *corridor*. This holds also in the regions of the highest orders, and the width of the photographic plate required for a lens or mirror of any given focus for recording the entire spectrum is measured by the width of the corridor. We have now the general rule that if we erect the corridor lines embracing the lines of longest wave-length which we are able to record, and which appear in the two spectral bands of lowest order recorded by the plate, when the superposed orders are separated by the auxiliary dispersion, all of the lines of shorter wave-length will be found between the two vertical lines which bound the corridor. This is true only if we are in a region in which the corridor is not blocked by the absence of an order, or orders, as called for by the theory of "absent orders." These regions which are very conspicuous in the bands of very high order should, of course, be avoided. The spectrum of iron, for the corridor embraced by the double line of wave-length 6137.5 is shown by Fig. 2A for the orders between 13 at the left and 14 at the right, and it is at these two points, the corridor lines are erected. The orders from 13 to 19 are shown in this figure, as well as two orders for neon (longer lines), showing the possibility of introducing a comparison spectrum.

We will now take up the experimental work. An echelette for this purpose must be superior in quality to those in general use, and I was able to find only two, made some time ago, which could be used, in the preliminary work, one of 1800 lines to the inch, which had the very unusual property of throwing a large percentage of monochromatic light, say the green mercury line, into a single order (the 18th), or by a very slight change in the angle of incidence into two adjacent orders, corresponding to the so-called "single- and double-order positions" of the Michelson echelon. The other grating, having only 1440 lines to the inch, concentrated most of the energy in the fourth-order spectrum on one side of the central image, but on the other side threw sufficient light into a number of orders in the region around the twenty-fifth to yield fairly strong spectra. A transmission replica of this grating, which was the one used in making the 18-inch mosaic grating for the 18-inch "Schmidt" telescope at Mt. Palomar which throws over 90 percent of the light into the *first*-order spectrum was employed instead of a prism for the separation of the superposed spectra of high order.

The preliminary experiments were made in the summer of 1946, with improvised apparatus, employing sunlight. The collimator was made from an old telescope of four-inches aperture and a focus of five feet, the eyepiece replaced by a very short slit. The light from this fell on the echelette and was reflected into the camera lens of 3-inches aperture and thirty-inch focus. The dispersing replica grating was mounted in front of the camera lens, with its lines perpendicular to those of the echelette. With the 1800 grating

crossed with the 1400 replica, I resolved solar lines of 0.3A separation, obtaining with a single exposure on a photographic film only two inches in length the complete solar spectrum from $H\alpha$ in the red to $H\gamma$ in the violet. The dispersion was equivalent to that given by a grating of 14,400 lines to the inch, requiring a plate nearly two feet in length, for its recording in the first-order spectrum. The most surprising thing was that the area of the echelette actually used was less than an inch square, only enough barely to separate the D lines of sodium in the first-order spectrum, 7A in comparison to 0.3A obtained in the high orders.

Using the 1440 echelette, with which one has the light fairly bright in the region between the 25th and 32nd orders, one can examine the field as we alter the angle of the camera. The $H\alpha$ line is found, as is to be expected, say at the extreme left in the lowest order, and again in the order above at the extreme right of the field, and as we move the camera towards the region of the higher order the $H\alpha$ line appears, over and over again, one order at a time in the ascending orders.

In the continuation of the work, the iron arc has been employed as a source, with a Cooke objective of 4-inches aperture and one meter focus for the camera, and the same collimator as in the earlier work. Figure 2A shows that the corridor is not blocked to any extent with the 1800 echelette when 486 is in the 17th order. The region of the group of superposed spectra in which a given photograph has been made will be "indexed" or labeled by the order number of the iron line 4860, which I shall call the "index line." This line has been chosen on account of the ease with which it can be identified, as it heads a group of three double lines and is bordered on the other side by a wide region in which only a few very faint lines are found. If the echelette is oriented in the position to yield spectra in which short wave-lengths are at the left, and the replica grating is oriented so as to give the red band at the bottom and the violet at the top, we must first find the order number of 4860 for a given position of the camera or telescope. (Most of the work has been done with 4860 in the 18th order.)

To do this the camera is moved into the posi-

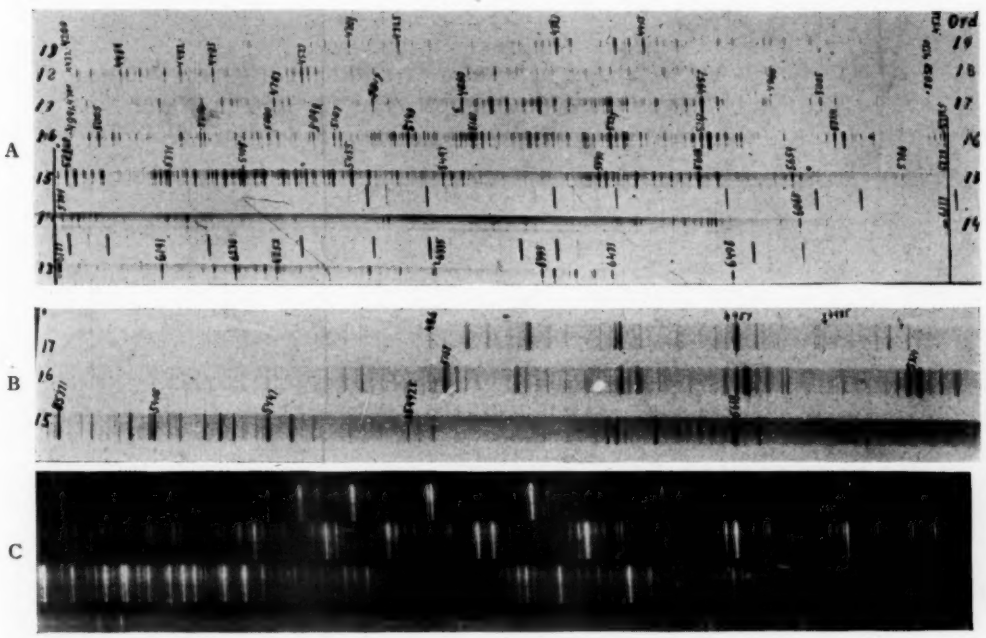

FIG. 2.

TABLE I.

Order No. for 4860	Next lower order coincidence				
6	5820	15	5208	23	5085
7	5760	16	5185	24	5071
8	5550	17	5163	25	5062
9	5470	18	5145	26	5054
10	5400	19	5130	27	5046
11	5345	20	5115	28	5040
12	5300	21	5105	29	5034
13	5265	22	5091	30	5027
14	5235				

tion in which the central image is at the center of the field and seen as a *vertical* spectrum of the very short slit, formed by the replica, the red being at the bottom and the violet at the top. We next move the camera or eyepiece until the very short first-order spectrum appears, slightly turned from the vertical with respect to that of the central image. We observe and count the successive orders as they come into view, descending in the field, and becoming more oblique as we proceed to higher orders. If difficulty is found in identifying the iron lines in the low orders, the green line of the mercury arc can be employed. We thus set the camera for the approximate position for the 18th order for the 4860 line. If we have hit upon the position giving 4860 in the 18th order, we shall find that it coincides with the position of $\lambda = 5146$ in the spectrum below it, for, as we have shown by Fig. 1, in this case, $(4860 \times 18)/17 =$ the wave-length of the point of coincidence of 4860 in the spectrum below. If this relation is not found to hold, we move the camera to the next order one way or other until we have 4860 in the 18th order. A similar relation holds for the coincidence point for all possible order numbers of 4860, and once having calculated these coincidence points, one can tell by a mere glance at the spectrogram, the orders in which he is working, higher orders being above and lower ones below the band in which 4860 is found. Obviously any other line could have been chosen for this purpose. As the order number of the 4860, or index, line passes into higher orders, its coincidence point in the band below moves progressively by steadily decreasing shifts to regions of shorter wave-lengths,

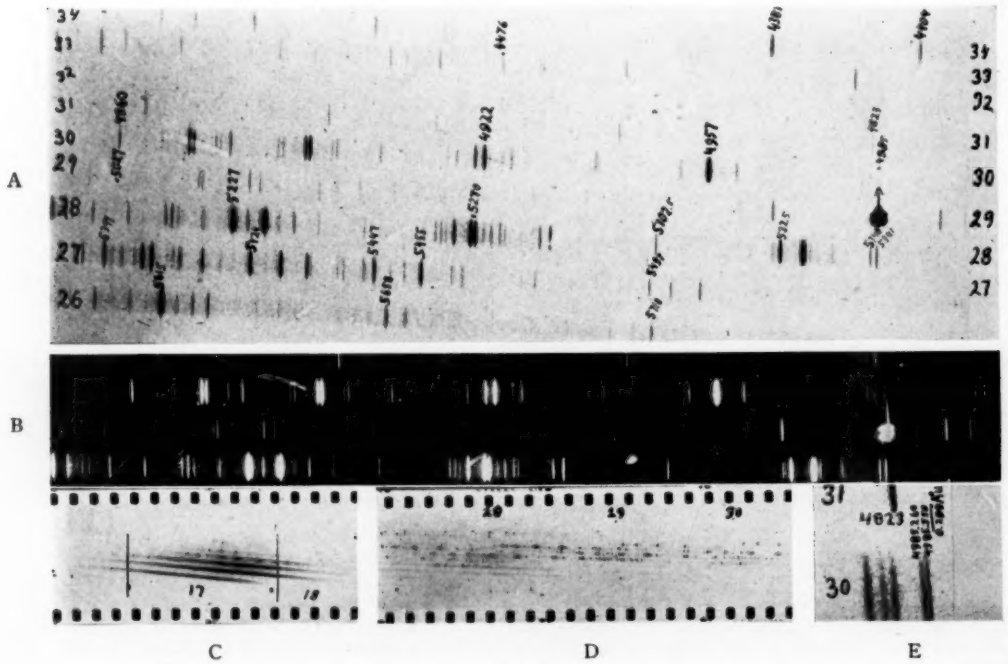

FIG. 3.

approaching 4860 as a limit, which would be reached practically for a very high order number, say $n = 10{,}000$. Figure 2B and C show the first shift to shorter wave-lengths of the coincidence point as the order of 486 increases from 17 to 18. The points of coincidences, expressed in wave-lengths, are given in Table I. This table, compared with a good map of the iron spectrum enables one to identify the order number of the spectrum band in which the 486 line is found, regardless of the number of lines to the inch in either the echelette or the dispersing replica, by merely observing the coincidence point.

Figure 2B shows the index line 486 in the 17th order in coincidence with 5163 in the 16th order, while Fig. 2C records the coincidence at 5145 in the 17th order when 486 is in the 18th. In Fig. 3A the index line in the 30th order coincides with 5027 in the 29th order. On Fig. 2A the "ghosts" are very strong, as the echelette was ruled at a time when a large periodic error had developed suddenly during the ruling. Figure 2B shows the same region (with slightly greater enlargement) made from a region of the echelette showing much weaker ghosts.

As has been stated previously the resolution of two solar lines with a separation of 0.3A was obtained in the earlier work. There are two iron lines having very nearly the same separation. These were shown clearly resolved in the 30th-order spectrum recorded on Fig. 3A. Their wave-lengths, as given in Harrison's tables, are 4985.559 and 4985.260 or a separation of 0.299A. The position of these lines is indicated by an arrow above the circular black spot at the right. As they are nearly invisible on the "negative" (or white print) a "positive" print of 3 orders is shown immediately below in which they are seen clearly resolved, while a print of a still higher enlargement is shown on Fig. 3E. These lines were resolved also in the 17th order of Fig. 2.

With the 1800-line grating a dispersion of 4.6 angstroms to the millimeter was obtained, while with an original echelette ruled on a silver film with 1440 lines to the inch, and the collimator and camera set at an angle of 90°, the dispersion was 3.1 angstroms per mm. At an angle of 30°, the usual angle employed, it was 4.3 angstroms.

On Fig. 3C a contact print from a contact positive of the original 35-mm film (Eastman Panatomic X) shows the concentration of the total light in the region in which the 486, or index line is in the 17th order. The "corridor" lines have been drawn on the print. A more "open" corridor is obtained when the grating is set for 18th-order concentration. Figure 3D is a similar contact print in the region between the 27th and 30th order from the film from which the enlargement (Fig. 3A) was made. This film was made with the 1440 echelette, in the region of rather weak concentration of light in several very high orders, most of the light going into the 4th and 5th orders on the other side of the central image, as has been alluded to previously. It is believed that further improvements can be made as soon as gratings especially designed for this work can be ruled.

Observation of the wide reflecting edges of the 1800 echelette with the "micro-interferometer" by W. E. Williams, showed that they were uniformly plane to within a very small fraction of a wave-length, but that a rough region existed along the top of each ridge. This we have previously expected from the performance of many of our gratings, and it is probably responsible for the scattered light which is very strong in the case of the aluminized 1800 echelette. What remains to be done is to get a high concentration of light in one or two of the higher orders, and to get rid of, or reduce, the faint or absent orders. Dr. Shane thinks this can be done by a special type of ruling, combined with a certain definite angle of incidence.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

A Tunable Narrow-Band Optical Filter

BRUCE H. BILLINGS*
Research Laboratories, Polaroid Corporation, Cambridge, Massachusetts
(Received July 18, 1947)

Polarization interference filters are described which can have a pass band ranging from a fraction of an angstrom to several hundred angstroms in width. The pass band can be shifted to any desired region of the spectrum. These tunable filters are based on the fixed filters discussed by Lyot and Evans. The transmission band is formed by the superposition of the polarized channel spectra, produced by x-cut plates of quartz or other birefringent media placed between parallel polarizers. The birefringent plates have thicknesses in the ratio 1:2:4 etc. The tuning is accomplished by changing the retardation of successive elements so that transmission maxima in the various channel spectra coincide at the desired wave-length. The retardation change can be made mechanically, for example, by stretching supplemental plastic sheets in series with the filter elements, or can be made electrically by using Kerr cells or crystals with high electro-optic coefficients, such as ammonium dihydrogen phosphate. The additional retardation never has to exceed a full wave at the wavelength of peak transmission. The measured transmission of an experimental filter is shown. The electrical tuning method is particularly adapted to cathode-ray oscillograph presentation of spectra. The filter also has possible application in color reproduction and colorimetry. With a pass band of a half-angstrom line of sight motion of solar prominences could be determined by the use of the Doppler shift of the prominence radiation.

IN 1933 Lyot[1] discussed a filter which was planned to have an aperture of 25 mm and a pass band in the green of only 1A. Using the technique described in his paper, various investigators[2] have built filters of wide apertures and narrow transmission bands. The filters of this type have been fairly expensive to make, and have been used so far in the study of solar prominences. An improvement in such a filter would be to make it variable so that the pass band could be shifted at will over a considerable range of wave-lengths. Such an improvement has been described by Lyot[3] and by Evans[4] but is difficult to manipulate.

In this paper a simple tuning method is described which makes it possible to set the pass band of the filter at any point over a considerable range.

Both the fixed and, more especially, the tunable filter represent new scientific tools which have many possibilities for exploitation. The filters can be made with any pass band from a fraction of an angstrom to hundreds of angstroms. They can also be made to function in the ultraviolet and in the infra-red through the quartz-transmission region. In colorimetry and color reproduction there are many applications for a filter of perhaps 100-angstroms band width which could be tuned in discrete jumps rapidly through the visible spectrum. With the high speed provided by electrical tuning methods it should be possible to make spectral analyses of very short-lived phenomena, such as the discharge of an Edgerton lamp. These spectra could be displayed on a cathode-ray oscillograph. Presentation on cathode-ray oscillographs of other absorption or emission spectra seems quite straightforward with electrical tuning methods. With very narrow band filters it should be possible to make an analysis of the fine structure of Raman lines. Another application for the tunable filter is for the analysis of three-dimensional motion of solar prominences. A more detailed description of the astronomical applications will be made in another paper.

The peak transmission of the filters will depend very much on the method of construction. With a many-element filter using sheet polarizers the peak transmission may be of the order of 10 percent. With quartz or calcite polarizers and a laminated construction, the peak transmission might perhaps be as high as 35 or 40 percent. The angular aperture of the filter will also depend on its construction and on the narrowness of the desired pass band. For filters of several hundred-

* Now with Baird Associates Inc. Cambridge, Massachusetts.

[1] B. Lyot, Comptes rendus **197**, 1593 (1933).
[2] Y. Ohman, Nature **141**, 291 (1938); J. Evans, Pub. Astro. Soc. Pacific **52**, 305 (1940).
[3] B. Lyot, Ann. Astrophys. **7**, No. 1–2, (1944).
[4] J. Evans, Cientia e Investigacione, to be published.

angstroms band width it might be of the order of 20 degrees total field, whereas for a one-angstrom band filter the angular field might be of the order of a degree.

THEORY OF THE FILTER

In the basic filter the narrow band is formed by the superposition of the polarized channel spectra produced by many retardation plates. An x-cut plate of quartz or other birefringent crystal, when placed between crossed or parallel polarizers, transmits light at periodic intervals through the spectrum. In a spectroscope such a plate shows a series of dark bands across the spectrum—hence, the term "channel spectra." In the pile of alternate plates and polarizers making the filter, the thickest plate is chosen so that the width between a pair of black bands is the width of the desired pass band. The second plate is chosen half the thickness of the first plate. This second plate will cancel out every other pass band of the second plate. This is continued until the final pass bands of the system are far enough apart to be removed by auxiliary dye filters or filters of other types.

The transmission of a four-component filter and the transmission of the individual elements is shown in Fig. 1. Associated with the pass bands in this figure are some low intensity secondary bands. The location and height of these secondary pass bands is shown by a mathematical analysis of the filter transmission.

For the transmission of an N plate filter we can write:

$$T=(A'\cos\theta\cos2\theta\cos4\theta\cdots\cos2^{N-1}\theta)^2, \quad (1)$$

where $\theta=\pi(n_e-n_0)d/\lambda$, d is the thickness of the thinnest plate, n_e and n_0 are the ordinary and extraordinary indices of the crystal, and λ is the wave-length. We can replace $\cos n\theta$ with its equivalent:

$$\cos n\theta=(e^{ni\theta}+e^{-ni\theta})/2. \quad (2)$$

This can be rewritten:

$$\cos n\theta=(e^{ni\theta}/2)(1+e^{-2ni\theta}). \quad (3)$$

If this substitution is made, Eq. (1) becomes:

$$T^{\frac{1}{2}}=A'\frac{e^{i\theta}}{2}(1+e^{-2i\theta})\frac{e^{2i\theta}}{2}(1+e^{-4i\theta})\frac{e^{4i\theta}}{2}(1+e^{-8i\theta})\cdots \times\frac{\exp[2^{N-1}i\theta]}{2}(1+\exp[-2^Ni\theta]). \quad (4)$$

This can be expanded to give:

$$T^{\frac{1}{2}}=(A'/2^N)e^{i\theta\sum_{N=1}^{N}2^{N-1}} \times[1+e^{-2i\theta}+e^{-4i\theta}+e^{-6i\theta}+\cdots e^{-i\theta\sum_{N=1}^{N}2^N}]. \quad (5)$$

Now:

$$\sum_{N=1}^{N}2^{N-1}=2^N-1;\quad \sum_{N=1}^{N}2^N=2^{N+1}-2, \quad (6)$$

$$T^{\frac{1}{2}}=\frac{A'}{2^N}(\exp[i\theta(2^N-1)])\frac{1-\exp[-2i\theta(2^N)]}{1-\exp[-2i\theta]}; \quad (7)$$

or:

$$T^{\frac{1}{2}}=\frac{A'}{2^N}\frac{\exp[2^Ni\theta]-\exp[-2^Ni\theta]}{2i}\cdot\frac{2i}{e^{i\theta}-e^{-i\theta}}, \quad (8)$$

which becomes:

$$T=[A'\sin2^N\theta/2^N\sin\theta]^2. \quad (9)$$

For comparison with this equation, consider the expression for the transmission of a diffraction grating:

$$T=A(\sin^2\alpha/\alpha^2)\cdot(\sin^2n\varphi/\sin^2\varphi), \quad (10)$$

where

$$\alpha=(\pi/\lambda)d\sin\theta,\qquad \varphi=(\pi/\lambda)(d+b)\sin\theta. \quad (11)$$

In these equations n is the number of lines in the grating, d is the width of the transparent line of the grating, b is the width of the opaque lines in the grating, and θ is the angle from the grating normal. It is apparent from these last

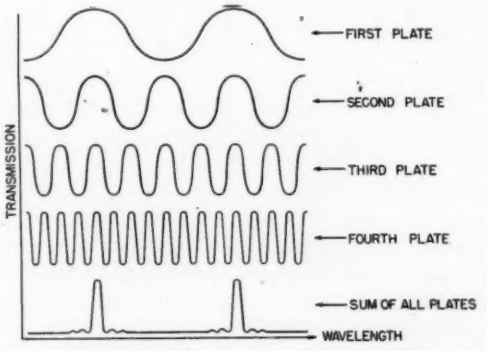

FIG. 1. Transmission between parallel polarizers of the individual retardation plates of a four-element filter and the transmission of the assembled filter.

equations that the N plate filter is analogous to a grating with 2^N lines, and the secondary pass bands are the equivalent of the secondary lines associated with the main lines of a diffraction grating.

ANALYSIS OF TUNING METHOD

If the retardation or thickness of the plates making up the pile were changed in such a way that the 2, 4, 8 ratio in retardation could be maintained, the tuning would be easily accomplished. Since some plates may have retardations of several thousand waves, the 2, 4, 8 ratio cannot be easily kept in a tuning system which uses variable retardation plates in series with the fixed elements. Most variable plates, such as Kerr calls, stretched sheets of plastic, or rotating birefringent plates have a retardation of only one or two waves. A scheme was devised by Lyot[3] in which the individual plates of the filter were cut into two wedges which could be slid across each other so as to vary the length of the light path through the combination. Since each plate has to be so tuned, the adjustment is very difficult and the filter is not considered variable.

It will be shown below that the 2, 4, 8 ratio can be considerably upset without introducing significant distortion of the filter pass band. In fact, the ratio can be departed from to such an extent that tuning can be accomplished by including with each element a low order variable retardation plate, such as a Kerr cell.

The retardation of the Kerr cell is a function of the voltage applied to it. Retardations up to full wave in the yellow are not too difficult to attain. The problem is to tune the filter and yet keep the change in retardation of the individual plates under a definite limit—in particular, under one half-wave.

The wave-length at which the filter transmission is to be peaked will be expressed in terms of the corresponding retardation of the thickest plate. The technique used in tuning will be to vary the retardation of each plate so that its nearest transmission peak is shifted to the point at which the over-all transmission is to be a maximum. In order to do this, the argument of the cosine for each term in the expression for the filter transmission must be a multiple of π at this point. We will operate in the region $p\pi<\theta<(p+1)\pi$, i.e., in the region where the first plate has a retardation Γ, where $p\lambda_1<\Gamma<(p+1)\lambda_2$ and λ_1 is the wave-length at which the angular retardation is $2p\pi$, and λ_2 the wave-length at which it is $2(p+1)\pi$. We will suppose that we wish to shift the pass band to a point where the angular retardation of the last plate is $2\pi(2^{N-1}p+x)$. At this point the argument of the cosine of the last plate can be written:

$$\theta_{Nx}=2^{N-1}\theta_{1x}=x\pi+2^{N-1}p\pi. \tag{12}$$

At the transmission peak of the Nth plate which is nearest to x:

$$\theta_{Nn}=(n_N+2^{N-1}p)\pi, \tag{13}$$

where n_N is an integer.

n_N is chosen so that $|x-n_N|$ is a minimum. If the transmission peaks of the Nth plate are numbered starting with the filter transmission peak as zero, n_N is the number of the peak nearest x.

We wish to multiply θ_{Nx} by a number "a" such that θ'_{Nx}, the value for the shifted plate, will equal θ_{Nn}.

$$\theta'_{Nx}=a\theta_{Nx}. \tag{14}$$

From this equation we find for a:

$$a=(2^{N-1}p+n_N)/(2^{N-1}p+x). \tag{15}$$

For the Kth plate:

$$a=[2^{N-1}p+(2^{N-K})n_K]/[2^{N-1}p+x], \tag{16}$$

where n_K is the number of the peak of the Kth plate nearest to the point at which maximum transmission is desired, and x refers again to the Nth plate.

Since the produce $(2^{N-K})n_K$ may be independent of K over quite a range of K, this same value of a_K holds for several plates. In this case we can write a single symbol, a_i, for a series of plates where i is the number of the thickest plate which has a_i as a factor:

$$a_i=a_K=a_{K+1}\cdots=a_i. \tag{17}$$

For the transmission of a filter tuned to a point x:

$$T'=A^2[\cos a_1\theta\cos 2a_2\theta\cos 4a_3\theta\cdots]^2. \tag{18}$$

Because of the equality of many of the a_K's, this can be rewritten in the form of Eq. (9):

$$T' = \prod_{i=1}^{p} (\sin 2^{n_i} a_i \theta / 2^{n_i} \sin a_i \theta)^2, \qquad (19)$$

where

$$\sum n_i = N. \qquad (20)$$

If, instead of shifting to the nearest peak of each plate, we shift to the next highest peak, the equations are unchanged in form but n_K becomes the next highest peak to $x/2^{N-i}$.

CONSTRUCTION AND CHARACTERISTICS OF A DEMONSTRATION FILTER

To demonstrate the properties of the filter, a four-element demonstration model was built by Mr. Edgar Barr. The thinnest element of this filter had a retardation of two waves in the green. Both the fixed and variable elements of the filter consisted of plates of plastic. For the fixed plates the plastic was polyvinyl alcohol which had been stretched to the desired orientation and laminated between glass plates with Acryloid F–10.[5] The variable plates consisted of sheets of polyvinyl butyrate plasticized with 15 percent butyl sebacate. These sheets were mounted in a clamp which could be adjusted with a screw so as to stretch the sheet to the desired retardation.

TABLE I.

Plate number	Retardation
1	8880 mu
2	4480 mu
3	2250 mu
4	1050 mu

Each element of the filter was mounted on an optical bench in such a way that it could be swung out of the system.

The measured retardation of the filter components is shown in Table I. With a slight adjustment of the tuning elements associated with these plates, their retardation is made to have the 2, 4, 8 ratio at which they have a single pass band.

To operate the filter, light from a ribbon-filament lamp was projected through the system and through a carbon disulfide prism onto a screen. All the plates except the thickest were removed from the system. On the screen there appeared the channel spectrum of this first plate. The screw on the corresponding variable sheet was then adjusted until a pass band of this plate corresponded with the desired transmission band of the whole filter. The second plate was then thrown into the system and the same adjustment made. When all the plates

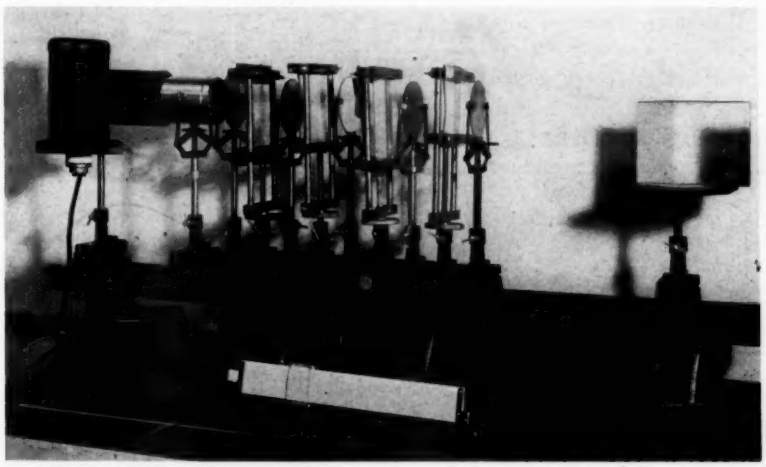

FIG. 2. Four-element plastic tunable filter. The circular filter elements are laminated to polarizers. The tuning elements are mounted with screw clamps. Examination of the behavior of the filter is made with the prism.

[5] This material is made by the Resinous Products Company in Philadelphia, Pennsylvania.

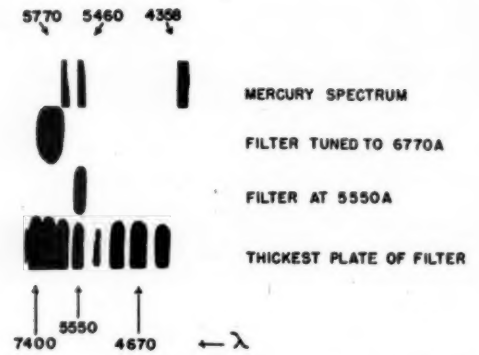

FIG. 3. Carbon disulfide prism spectrum of the light transmitted by an untuned four-element filter and spectrum of the light transmitted by the same filter tuned to a different wave-length region. The source is a tungsten lamp. A mercury arc spectrum is shown for comparison.

had been adjusted, a single band was transmitted by the filter. The complete filter is shown in Fig. 2.

In Fig. 3 is shown a series of photographs of the filter spectra made with the carbon disulfide prism. In the top line is the spectrum of the thickest plate between parallel polarizers. In the next line is the green line transmitted by the completely assembled filter. The next line shows the filter tuned to one of the red lines transmitted by the thickest plate. A mercury arc spectrum is shown for comparison.

It is apparent that the filter can only be tuned over a range which is equivalent to the separation between two successive pass bands in the fixed filter. As the filter pass band is shifted farther from the fixed-filter position, the 2, 4, 8 ratio of retardation in the individual elements is more and more upset. The experiments in the preceding section seem to indicate that this upsetting of the 2, 4, 8 retardation ratio does not give rise to any extraneous pass bands. To test this further a series of transmission curves has been plotted for a five-element filter whose first four elements are identical with the experimental filter described above. The transmission of the filter with no retardation in the variable plates is given in Fig. 4.

To go to the first peak of the thickest plate:

$$x=1. \tag{21}$$

In Eq. (16):

$$p=2, \quad n_K=0, \quad a_1=32/33. \tag{22}$$

Since $2^4\times(1/32)<1$, the same a may be used all the way through to the last plate for which $a_N=1$. The transmission is given by the equation:

$$T_1=[\cos a\theta\cos 2a\theta\cos 4a\theta\cos 8a\theta\cos 16\theta]^2. \tag{23}$$

At the second peak $a_1=16/17$, but only the first three plates need to be altered. This transmission is plotted in Fig. 5. To go to the third peak requires two a's: $a_1=32/35$ for the first three plates and $a_2=34/35$ for the fourth plate. The transmission is given by the equation:

$$T_3=[\cos a_1\theta\cos 2a_1\theta\cos 4a_1\theta\cos 8a_2\theta\cos 16\theta]^2. \tag{24}$$

This transmission is plotted in Fig. 6.

In Fig. 7 is shown the transmission for the shift to the fourth peak.

It is apparent that over the range shown the filters have not degenerated as far as band width and signal-to-noise are concerned.

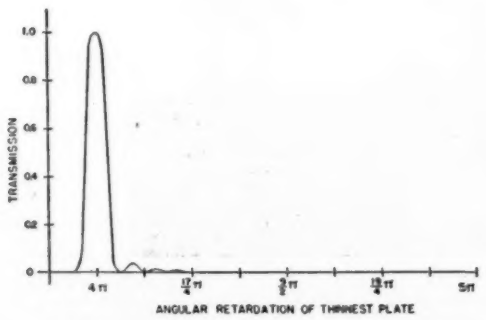

FIG. 4. Calculated transmission of a five-element filter with two waves retardation for the thinnest element.

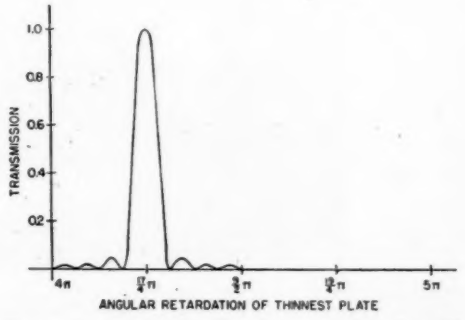

FIG. 5. Calculated transmission of a five-element filter tuned to transmit a wave-length at which the angular retardation of the thinnest plate is $17\pi/4$.

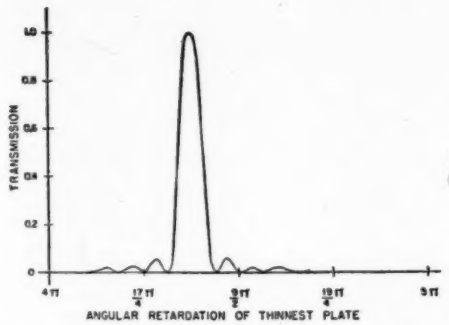

FIG. 6. Calculated transmission of a five-element filter tuned to transmit a wave-length at which the angular retardation of the thinnest plate is $35\pi/8$.

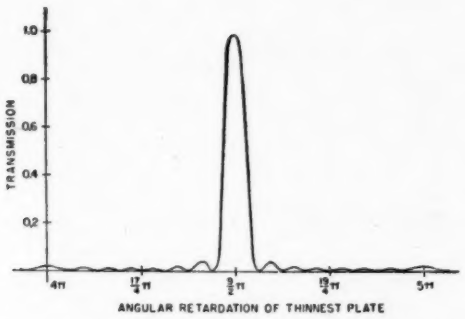

FIG. 7. Calculated transmission of a five-element filter tuned to transmit a wave-length at which the angular retardation of the thinnest plate is shifted to $9\pi/2$.

The a_i calculated on the basis of a maximum shift to the next highest n_K, gives more distortion to the undisturbed filter curve than those arising from a shift to the nearest n_K. The best factor will be the one which gives the longest wave-length interval with no secondary peaks.

To compare the effect of shifting to the nearest peak with a shift to the next highest peak, the peak at 2π was shifted up to $11\pi/4$. The equation for the transmission can be written:

$$T_{12}=[\cos(8/11)\theta\cos(20/11)\theta \cos4\theta\cos8\theta\cos16\theta]^2. \quad (25)$$

In this curve (see Fig. 8) there is a strong secondary maximum at $9\pi/4$. In Fig. 9 are shown the results of shifting the peak at 3π down to $11\pi/4$. The equation for the transmission is:

$$T_{12}=[\cos(12/11)\theta\ (\cos24/11)\theta \cos4\theta\cos8\theta\cos16\theta]. \quad (26)$$

The secondary maximum at $9\pi/4$ is tremendously reduced, and the effective usefulness of this filter is increased by a total range of $\pi/2$.

If, instead of tuning the filter by adding an ordinary retardation plate, the filter is tuned by adding achromatic phase plates, there is no change in the secondary maxima even at wave-lengths far removed from the point to which the filter is tuned. The transmission of a single element is:

$$T=A\cos^2a\theta. \quad (27)$$

With the addition of an achromatic plate this becomes:

$$T=A\cos^2(a\theta+\delta), \quad (28)$$

where δ is independent of θ. Tuning in this case does not change the spacing of the peaks in the curves but merely slides them along the wave-length scale.

In the non-achromatic case which has just

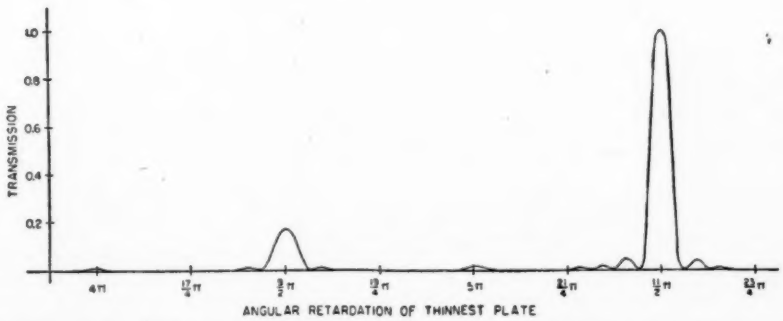

FIG. 8. Calculated transmission of a five-element filter tuned to transmit a wave-length at which the angular retardation of the thinnest plate is $11\pi/2$. In the calculation retardation was subtracted from each plate.

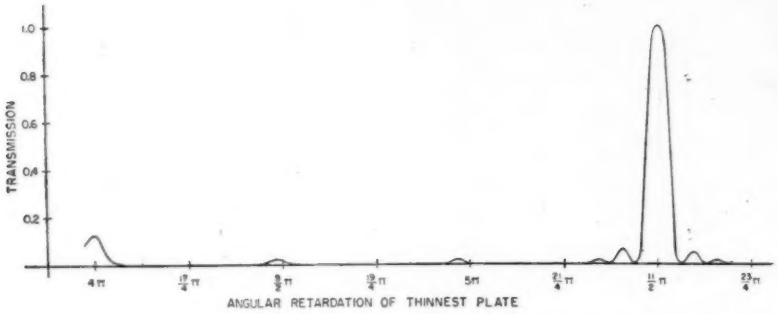

FIG. 9. Calculated transmission of a five-element filter tuned to transmit a wave-length at which the angular retardation of the thinnest plate is $11\pi/2$. In this calculation retardation was added to each plate.

been treated, the spacing between the peaks changes and it is the change in separation which changes the character of the secondary spectra. The maximum change in peak separation occurs when the θ is small, e.g., for a two-wave retardation plate. Since we have shown that the secondary maxima are not particularly distorting with a two-wave plate, it is apparent that any filter whose first plate has higher retardation than two waves can be successfully tuned over an interval of 2π in angular retardation.

In many of the applications of the tunable filter it is desirable to know how much retardation must be added to each plate to shift the pass band by a given amount. This information is especially necessary in building filters where the variable plates are adjusted by a single control. The retardation to be added to the Kth plate to produce a given shift can be written:

$$\Delta\delta_K = 2(\theta_{K0} - \theta'_{K0}), \tag{29}$$

where $2\theta_{K0}$ is the retardation of the Kth plate at λ_0. $2\theta'_{K0}$ is the retardation of the Kth plate at λ_0 after the transmission peak has been shifted. λ_0 is a reference wave-length. By substituting the value of θ'_{K0}, one gets

$$\Delta\delta_K = 2^K(a_K - 1)\theta_{1,0}. \tag{30}$$

From Eqs. (16) and (30) one finds

$$\Delta\delta_K = (2^N n_K - 2^K x/2^{N-1}p + x)\theta_{1,0}. \tag{31}$$

The formula for x can be written:

$$x = 2^{N-1}p(\lambda_0 - \lambda)/\lambda, \tag{32}$$

where λ is the wave-length at which peak transmission is desired. Substitution of Eq. (32) into Eq. (31) gives:

$$\Delta\delta_K = \frac{\lambda(n_K + 2^{K-1}p) - 2^{K-1}p\lambda_0}{p\lambda_0} \cdot 2\pi, \tag{33}$$

where this retardation is added in the region:

$$(2^K p\lambda_0/2n_K + 1 + 2^K p) < \lambda < (2^K p\lambda_0/2n_K - 1 + 2^K p). \tag{34}$$

For a filter which is to cover a continuous range between angular retardation $2(p-1)\pi$ and $2(p+1)\pi$, one can plot the retardation change for a given plate as a function of wave-length. This plot is shown in Fig. 10 for the first three plates of the filter described above.

TYPES OF TUNING AND APPLICATIONS OF THE FILTER

In the first part of this paper it was suggested that the variable retardation plates be Kerr

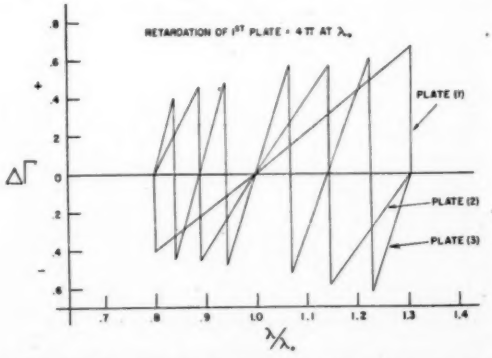

FIG. 10. Angular retardation in fractions of waves at λ_0 which must be added to the plates of a three-element filter to shift its pass band continuously from 0.8 λ_0 to 1.33 λ_0.

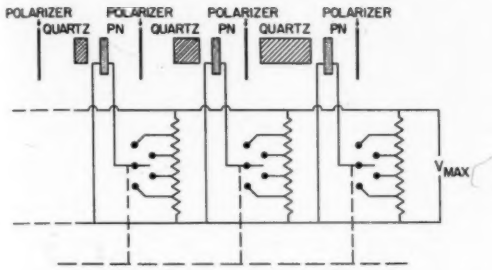

FIG. 11. Schematic diagram of a four-element tunable filter with an associated electrical network used to set the pass band to four discrete wave-lengths.

cells. The Kerr cell might be used rather loosely to indicate any material whose birefringence changes in an electric field. A material which seems to be especially adapted to the construction of these filters is crystalline $NH_4H_2PO_4$ or PN.[6] This uniaxial crystal has a high electro-optic effect and in particular it exhibits a strong longitudinal effect. For fields applied parallel to the direction of light propagation and at the same time parallel to the optical axis, there is a strong induced birefringence. In this crystal the effect is linear with field strength; and, since the light path and field strength are in the same direction, the effect becomes linear with voltage and independent of thickness. Because of this linearity a resistance network can be easily designed to give the appropriate voltages and, hence, retardation to the PN tuning element in a filter.

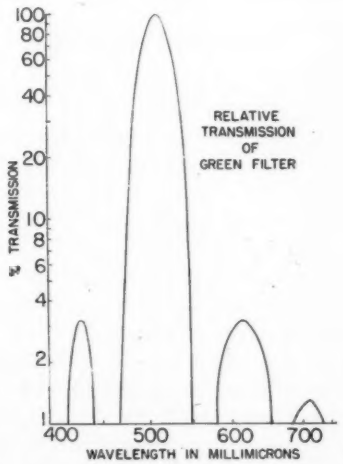

FIG. 12. Green transmission of a three-element tunable filter to be used for transmitting red, green, and blue for color television.

[6] This crystal is grown by the Brush Development Company in Cleveland, Ohio.

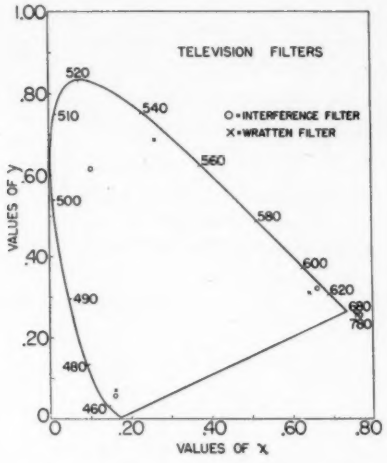

FIG. 13. Color of three possible positions for the pass band of a three-element tunable filter which might be used in color television. The Wratten *A*, *B*, and *C* are shown for comparison.

Figure 11 is the schematic diagram of a filter designed to be set to four discrete wave-lengths.

The electrical tuning method lends itself particularly to continuous scanning of a given spectral region. The curves of Fig. 10 can be considered as a set of sawtooth-voltage waves to be applied to the successive tuning elements of a filter. The basic cycle could be used as the sweep on a cathode-ray oscillograph, and the output of a photo-cell would be presented on the screen. This would give an instantaneous spectral curve of the light shining on the filter. Such cathode-ray tube spectral analysis might be useful for quick investigation of series of dyes or for process control which involved spectral-transmission curves.

The electrical tuning method is also interesting because of its possible application to sequential color television. The first three elements of the filter described in the previous section might be used in such an application.

By applying suitable voltages to the tuning

elements the filter can be made to pass either red, green, or blue light. The transmission curve of the green band is shown in Fig. 12. In Fig. 13 is shown the position of the three pass bands on a color triangle for equal energy.

The author is particularly indebted to Professor Donald H. Menzel and to Dr. E. H. Land for many stimulating discussions, and to the Polaroid Corporation in whose laboratory most of the experimental work was carried out.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Infra-Red Spectrometry from 15 to 25μ

EARLE K. PLYLER
National Bureau of Standards, Washington, D. C.
(Received July 18, 1947)

On account of the relatively small amount of energy emitted from sources in the wave-length region beyond 15μ, stray energy of short wave-length from the source becomes a considerable proportion of the energy indicated by the detecting system. By use of a reflection filter of LiF, this stray radiation is greatly reduced.

The infra-red spectra of cyclopentane, methylcyclopentane, ethylcyclopentane, and isopropylcyclopentane have been measured and each has a characteristic spectrum, which can be used in analysis.

Also, the spectra of cyclohexane, methylcyclohexane, ethylcyclohexane, and secondary butylcyclohexane have been measured and a number of bands which are characteristic of these substances have been observed.

INTRODUCTION

ALTHOUGH prisms of potassium bromide have been available for many years there have been comparatively few studies made of the absorption spectra of materials beyond the NaCl region in the infra-red. Some substances do not have bands in the spectral region between 15 and 25μ and the vibrations arising from C−H, O−H, C−O and other atomic groups within the molecules fall within the NaCl region.

The measurement of the absorption spectrum of substances in the longer wave-length region is more difficult than measurements in the near infra-red on account of the small amount of radiant energy given out by the source. The amount of radiation in the near infra-red region is much greater than the energy in the region beyond 15μ and if only a few percent stray radiation is present, it may make up a large fraction of the total energy present in the region from 20 to 25μ. On account of the small energy in comparison with the stray radiation, the determination of percent absorption of a substance with good accuracy is difficult. In the next section there are certain suggestions in regard to the experimental arrangements which may make this region more easily studied.

EXPERIMENTAL METHOD

In the region of 25μ slit widths of 1 mm were used and at 15μ the slits were 0.2 mm wide. By means of gears the slits were gradually opened. The rate of slit opening was insufficient to compensate for the diminution of energy. The deflections decreased as the wave-length increased and were about one-third as large at 25μ as at 15μ. The infra-red spectrometer was a Perkin-Elmer model 12A and a G.M. amplifier was used. In Fig. 1 is shown a record of the deflections in the region from 21 to 24.5μ. The records were obtained by advancing from the short wave-length region to the long wave-length region. Three shutters were placed in the beam. The first was opaque, the second was LiF and the third was CaF_2. The differences between these deflections indicate the amount of stray radiation at that wave-length. At 25μ the stray radiation was about 80 percent of the total

deflection. From the record in Fig. 1, it is seen that there is a difference in the zero position with the LiF and CaF_2 shutter. This difference is caused by the stray radiation from the region of 6.5 to 11μ. These shutters were about 4 mm in thickness. The percent of the stray radiation in the 25μ region can be decreased by lowering the temperature of the source. This, however, causes a decrease in the deflection.

In the measurement of the absorption bands of substances the amount of stray radiation passing through the cell will vary with the position and intensity of the absorption bands in the near infra-red region. It is necessary to take a zero reading with the fluoride shutter for the cell in the beam and for the cell out of the beam. When the amount of stray radiation is large, the correction is not sufficient on account of the loss in radiation by reflection of the shutter and the cell windows.

The experimental work would be simplified and better accuracy obtained if the stray radiation could be removed. Several methods have been tried such as use of transmission filters. The reflection of LiF has been found to be very satisfactory in the region from 15μ to 25μ. A slab of LiF was cut the same size as the plane mirror in the source box of the Perkin-Elmer instrument and was substituted for the aluminum mirror. The surface was polished so that the image of the globar source could easily be focused on the first slit of the instrument. The

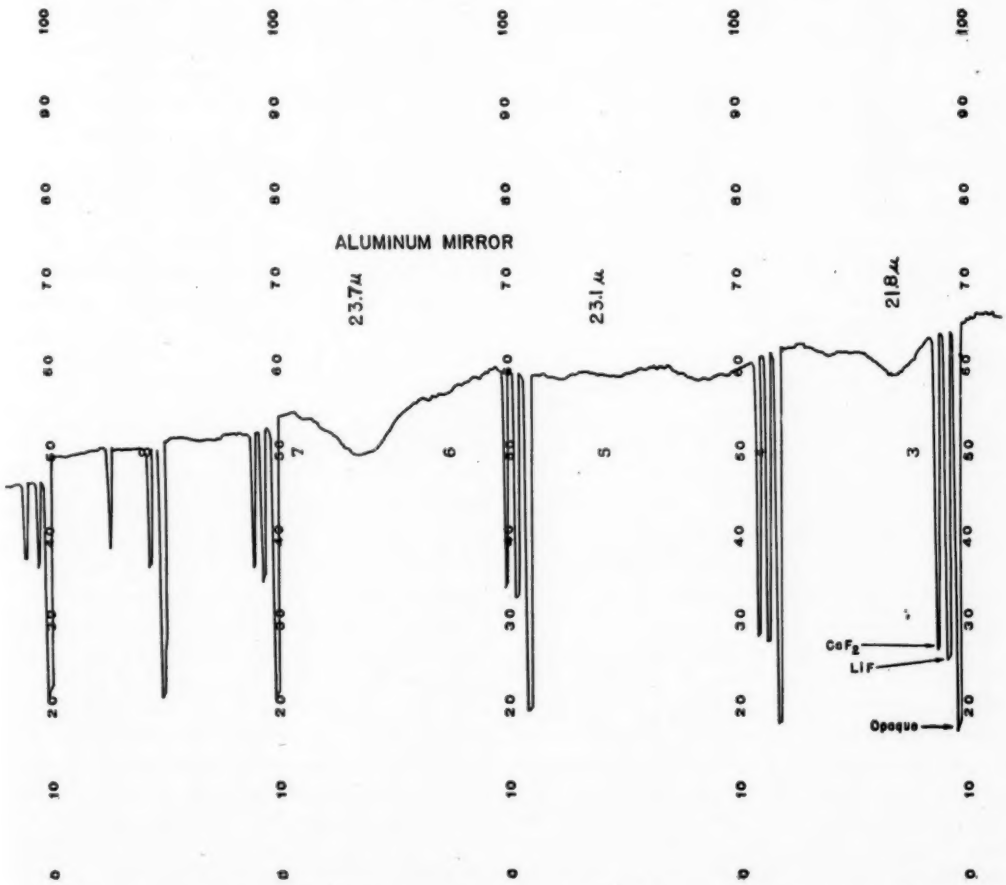

FIG. 1. The effect on the radiant energy between 21 and 25μ caused by interposing an opaque shutter and shutters of LiF and CaF_2.

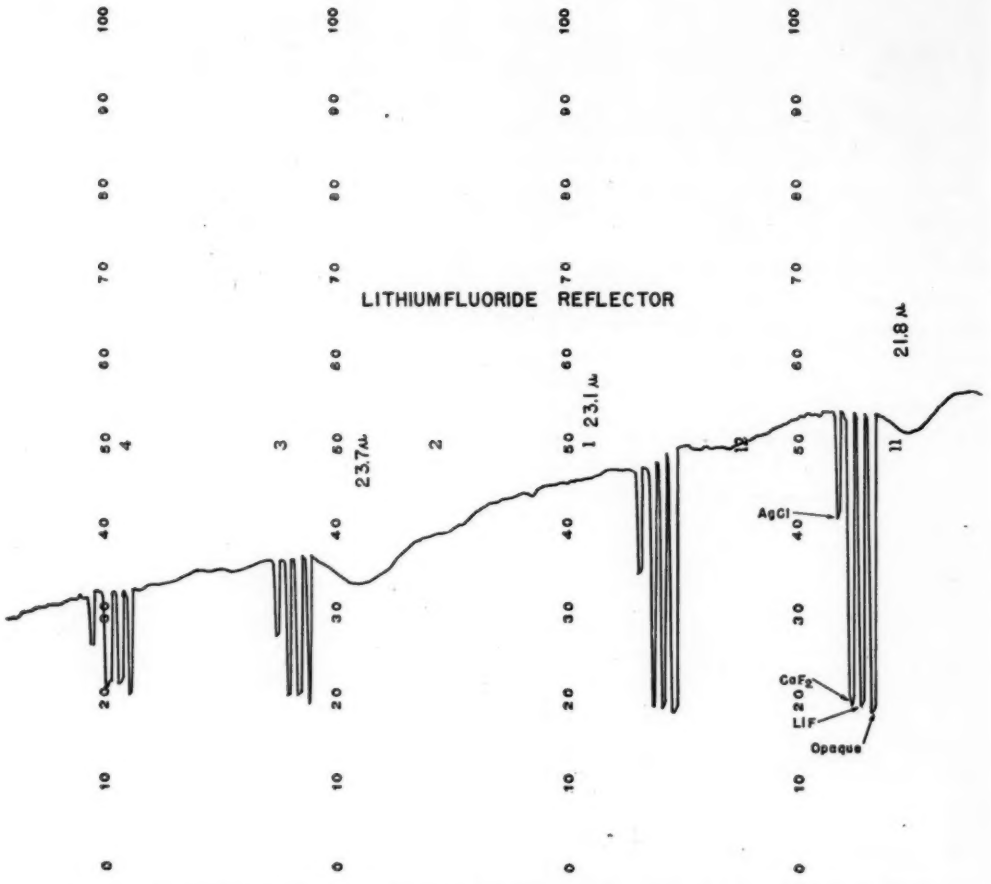

FIG. 2. A plate of LiF has been used instead of the aluminum mirror. The other conditions the same as in Fig. 1, except a shutter of AgCl has been added.

filter would be more efficient if the LiF plate was not highly polished. The short wave-length radiation would be diffusely reflected. This has been tried and further decrease in the stray radiation resulted.

In Fig. 2 is shown the record obtained with the LiF filter in the region from 21 to 24.5μ. The zero positions for the shutters are in the same order as the record shown in Fig. 1. That is, the opaque shutter was inserted first, the LiF second, and the CaF_2 third. The fourth deflection was caused by inserting a plate of silver chloride about 5 mm in thickness in the beam. The percent transmission of silver chloride for this spectral region is greater than that of sodium chloride but less than that of potassium bromide. By not allowing the short wave-length radiation to enter the instrument, a steadier condition is obtained, and there is much less zero drift.

There is some loss in energy by use of the LiF reflection filter. In Fig. 3 is shown the percent reflection of LiF and CaF_2 as compared to the aluminum mirror taken at 100 percent. The aluminum mirror reflects over 95 percent of the incident radiation in this region of the spectrum and these values of reflection for LiF and CaF_2 are approximately their percent reflection. The CaF_2 has its most intense maximum of reflection at about 32μ, but there appears to be a smaller maximum at about 23μ. The residual rays of

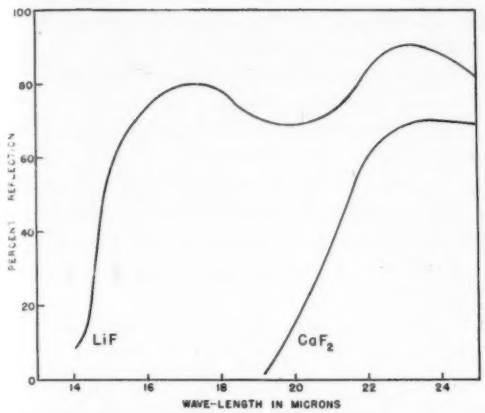

FIG. 3. The reflection of LiF and CaF_2 in the infra-red. The percent reflection is relative to the reflection of an aluminum mirror.

these materials and many other crystals have been measured by several investigators.[1]

The small loss in energy can be overcome by running the source at a higher temperature and increasing the gain of the amplifier. An increase in the energy is the greatest need for the accurate measurement of absorption spectra in the 25μ region. It has been mentioned that more sensitive detectors will be of considerable assistance. An attempt has been made to increase the energy by using different sources. A source containing 70 percent low resistance ceramic material has been made. The source was 50 mm long and 3 mm in diameter, and operated at 125 volts and 0.5 ampere. The resistance at room temperature was so high that it was necessary to heat the rod in order to pass appreciable current through it. It is well known that thorium oxide has a low emissivity in the region from 1.5 to 3.5μ and a high emissivity beyond 7μ. The radiation from the thorium oxide emitter was compared with the radiation from the globar. The intensity in the region between 2 and 3μ was made equal for the two sources by adjusting the temperature, and the region beyond 15μ compared. It was found that from 15μ to 20μ the thorium oxide emitter gave out more energy. At 15μ the energy of the thor-

TABLE I. Location of infra-red absorption bands.

Material	Wave-length (microns)	Wave-number (cm^{-1})
Cyclopentane (vapor)	18.31	546
Cyclopentane (liquid)	18.29	547
Cyclohexane (vapor)	19.12	523
Cyclohexane (liquid)	18.98	527
Isopropylcyclopentane (liquid)	18.19	550
	19.88	503
Methylcyclopentane (liquid)	18.66	536
	23.22	431
Ethylcyclopentane (liquid)	18.31	546
	23.41	427
Methylcyclohexane (liquid)	16.48	607
	18.29	547
	22.60	442
Ethylcyclohexane (liquid)	16.52	605
	17.30	578
	18.49	541
	20.00	500
	22.56	443
	23.39	427
Sec-butylcyclohexane (liquid)	16.29	614
	17.31	578
	17.62	567
	18.81	532
	19.77	506
	21.19	472
	21.95	455
	23.11	433

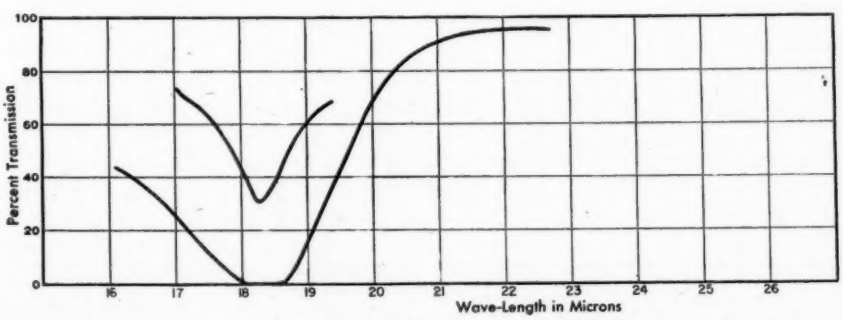

FIG. 4. The infra-red absorption band of liquid cyclopentane. Cell thicknesses were 0.2 mm and 1.25 mm.

[1] C. Schaefer and F. Matossi, *Das Ultrarote Spectrum* (Julius Springer Verlag, Berlin, 1930).

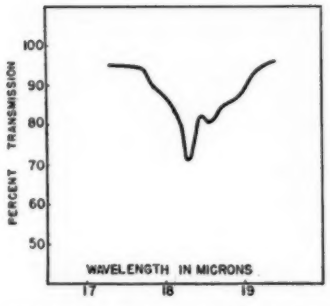

FIG. 5. The infra-red absorption band of cyclopentane vapor at 18.31μ for a cell of 140 cm, saturated at room temperature.

ium oxide was about 25 percent greater than the globar, and at 20μ the energy was about 5 percent larger. At 25μ the energy from the two sources was about equal. Another emitter containing a greater percent of thorium oxide is being made. However, if the energy at 25μ does not increase over the amount obtained from the globar, the gain in the 15μ to 20μ region is not sufficient to justify its use.

The instrument was calibrated by the use of several absorption minima in methyl alcohol vapor. These bands had been measured previously by Borden and Barker.[2] Also, atmospheric absorption of water vapor and carbon dioxide gave additional check to the calibration. The two water lines at 23.506 and 23.635μ were resolved by the instrument by using slits of 0.5 mm width.

Beyond 24μ the absorption bands are not easily observed. The deflection decreases very gradually in the region of the absorption band and the lowest point or region of greatest absorption is not readily apparent. Part of the broadening of the band is due to the width of slits. The first slit, if large, causes a mixing of many wave-lengths. This effect has been discussed by Schuster[3] and he defines the mixing of various wave-lengths over the second slit as

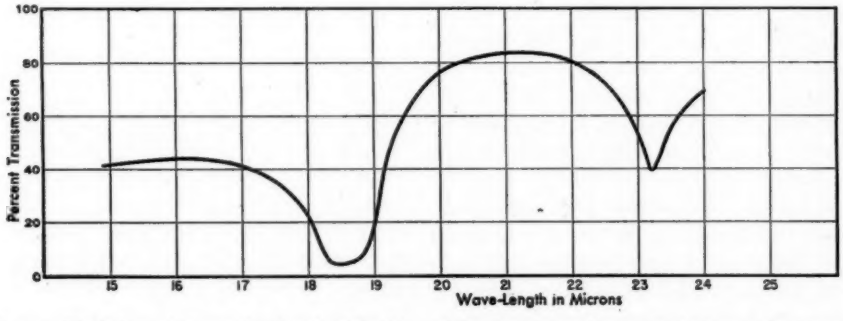

FIG. 6. The infra-red absorption bands of liquid methylcyclopentane for a cell thickness of 1.25 mm.

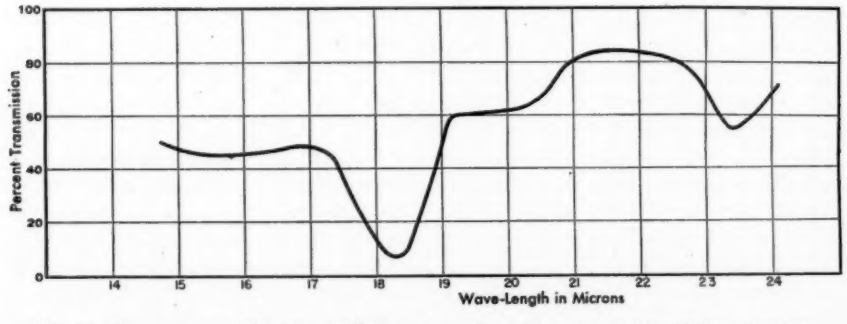

FIG. 7. The spectrum of liquid ethylcyclopentane showing bands at 18.31, 20.0 and 23.41μ.

[2] A. Borden and E. F. Barker, J. Chem. Phys. 6, 61 (1938).
[3] Sir Arthur Schuster, Astrophys. J. 21, 197 (1921).

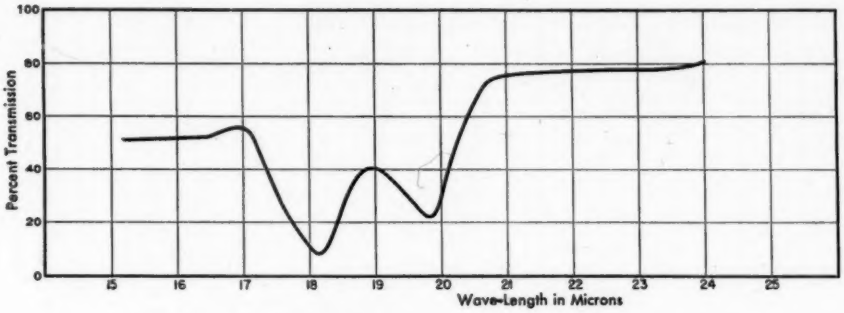

FIG. 8. The absorption spectrum of liquid isopropylcyclopentane.

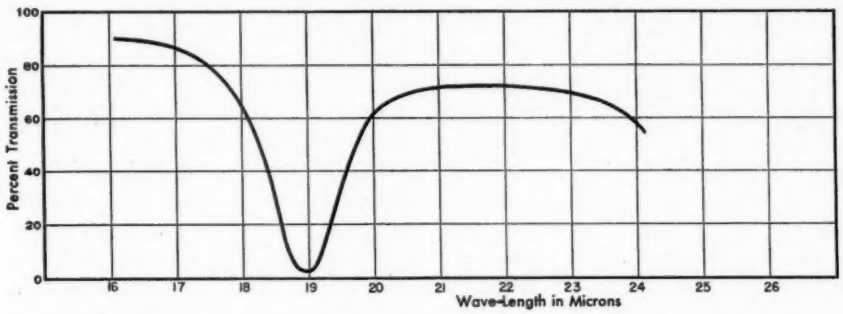

FIG. 9. The infra-red absorption spectrum of liquid cyclohexane for a cell 1.25 mm in thickness.

the purity factor. Schuster gives a table for the effect of first slit on the resolving power. It is a function of $f\lambda/D$ and is independent of the prism or exit slit, where f is the focal length and D the diameter of the mirror and λ is the wavelength of the radiation.

In the region of 25μ the KBr prism absorbs some of the energy. For this region a prism of an angle less than sixty degrees might be desirable. When the radiant energy is relatively intense, as in the near infra-red region, prisms of refracting angles greater than sixty degrees would give higher resolution. The seventy-two degree LiF prism, as used in a Perkin-Elmer infra-red spectrometer, is very satisfactory for high resolution in the near infra-red. In the region beyond 15μ the energy from the source is small and it is necessary to use slits up to 1.0 mm in width. There would likely be very little gain in resolution by using prisms of angles greater than sixty degrees in the 15 to 25μ region.

A more sensitive thermocouple of the type developed by Perkin-Elmer Corporation or a Golay detector, made by the Eppley Laboratory, would give larger output energy for the infra-red radiation and make the small energies more easily measured. By increasing the length of the slit of the spectrometer, a gain in the deflection can be obtained.

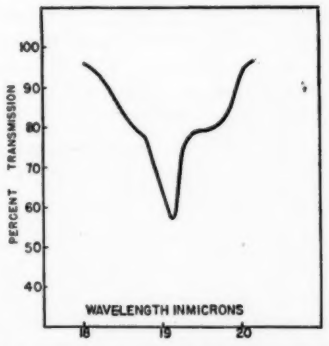

FIG. 10. The infra-red absorption band of cyclohexane in the vapor state at 19.12μ for a cell of 140 cm, filled with saturated vapor at room temperature.

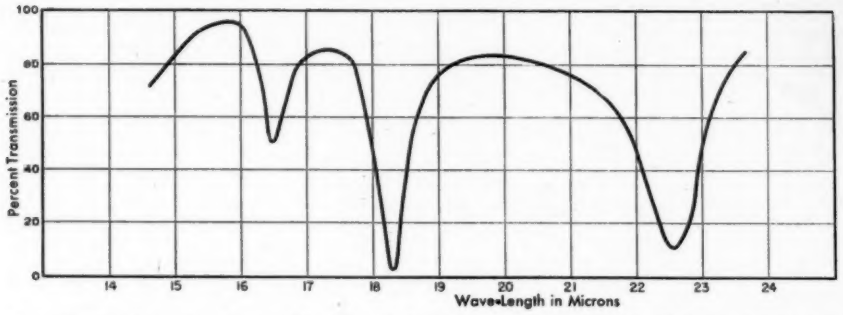

FIG. 11. The absorption bands of liquid methylcyclohexane for a cell with thickness of 1.25 mm.

THE ABSORPTION OF CYCLOPENTANES AND CYCLOHEXANES

The infra-red spectrometer, using the LiF reflection filter, has been used to measure the absorption spectra of 4 cyclopentanes and 4 cyclohexanes in the liquid state and cyclopentane and cyclohexane in the vapor state. The absorption spectra of cyclopentane and several of its substituents have already been measured in the infra-red region from 2 to 15μ. These substances have many bands in this region and there should be other bands in the region beyond 15μ which are characteristic of the skeletal vibrations.

In Fig. 4 is shown the absorption of cyclopentane for a cell of 1.25 mm, and 0.2 mm. The band at 18.29μ is of average intensity. There is considerable general absorption in the region from 15 to 17μ but with the 0.2 mm cell no band could be observed. Cells of 5 mm showed some general absorption in the region of 24 to 25μ, but no definite band was observed.

The saturated vapor in air was measured in this region with a path length of 140 cm. The room temperature was 72°F, and the liquid was placed in the chamber several hours before measurement. The only band which was observed occurs at 18.31μ and it consists of P, Q, and R branches. The interval in wave numbers between the P and R branches is about 20 cm^{-1}. The results are shown in Fig. 5.

The absorption spectra of methylcyclopentane, ethylcyclopentane, and isopropylcyclopentane have also been measured in this region. The cell thickness was 1.25 mm for all of these substances. In Fig. 6 is shown the absorption of methylcyclopentane. The band in the region of 18.29μ for cyclopentane is shifted to longer wave-lengths in methylcyclopentane. There is an additional band present at 23.22μ. All the cyclopentanes which have been measured in this study have considerable general absorption in the region from 15 to 17μ. This may, in part, be caused by the 18.29μ being quite diffuse. In addition to the characteristic cyclopentane band, there are to additional bands at 20.0μ and 23.41μ observed in ethylcyclopentane. The first of these bands is too diffuse for precise measurement. The absorption spectrum is shown in Fig. 7.

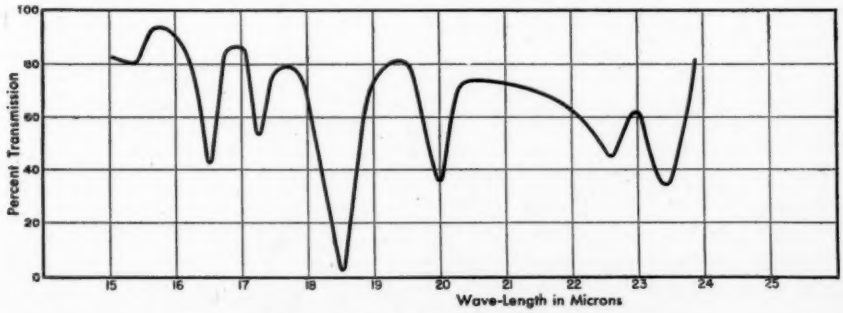

FIG. 12. The absorption bands of liquid ethylcyclohexane.

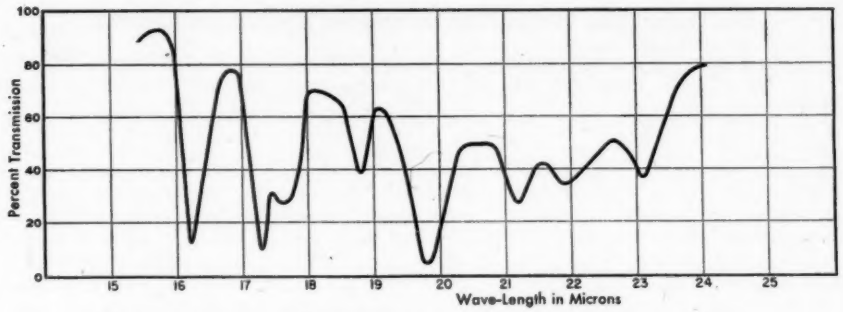

FIG. 13. The infra-red absorption spectrum of liquid sec-butylcyclohexane.

In Fig. 8 the percent transmission of isopropylcyclopentane is shown. There are bands at 18.19 and 19.88μ.

In Fig. 9 is shown the absorption of cyclohexane, the cell thickness being 1.25 mm for all the liquids. This hydrocarbon shows very little absorption in the region from 16 to 17 microns, and the band located at 18.98μ is sharper than the band observed in cyclopentane. In the region from 20 to 23μ, there is considerable general absorption and the absorption is increasing to 24μ. There is an indication of a band at 25μ, but the present experimental arrangements would not permit accurate measurements in this region. The absorption spectrum of the vapor of cyclohexane was measured under the same experimental conditions as used in measurement of the vapor of cyclopentane. A band was observed at 19.12μ, Fig. 10, and it consists of P, Q, and R branches. The band is not intense but somewhat stronger than the one observed in cyclopentane vapor. The region of general absorption from 20 to 23μ was examined, but no band was found for the vapor in the region of 21.8μ. The possibility of a band in this region was previously mentioned.[4] However, the repeating of the study of this region failed to show such a band. The low absorption of the cyclohexane vapor from 15 to 25μ makes long cells necessary for detecting any additional bands which may be present in this region.

The absorption spectrum of methylcyclohexane is shown in Fig. 11. A small band is observed at 16.48μ and the band of cyclohexane at 18.98μ is shifted to 18.29 and is very narrow. Also, an intense band occurs at 22.60μ. The bands observed in the cyclohexanes are usually sharp.

In Fig. 12 is shown the absorption bands of ethylcyclohexane. There are six bands in the region from 16 to 24μ. These bands are included in a region of about 200 cm^{-1} in width. This number of bands is greater than those observed in an equal spectral width in the 2 to 15μ region. In Fig. 13 is shown the absorption spectra of sec-butylcyclohexane. For this compound there are eight bands observed in the region from 16 to 24μ. The strong band at 18.98μ for cyclohexane is quite reduced in intensity and is observed at 18.81μ. The observed bands are listed in Table I and it is estimated that the positions are accurate to 0.04μ.

In making these observations it was found that the experimental difficulties are no greater from 15 to 23.5μ than in the region 2 to 15μ. To extend the work to 25μ, several experimental changes are necessary. Whether this region of the spectrum will be of value in analysis depends on the materials studied. If no suitable bands are found in the sodium chloride region, it would be of value to investigate the region beyond 15 microns. The cyclopentanes and cyclohexanes are weak in absorption beyond 15μ. However, in the near infra-red there are some very intense bands. Measurements have been made of the spectra of a number of other substances such as benzene, styrene, toluene, xylenes, and other hydrocarbons, and in some cases a cell of 0.05 mm shows absorption of over 80 percent. It seems that the region beyond 15μ is well suited to the study of some compounds, and that it will be useful in analysis.

[4] Earle K. Plyler, Ralph Stair, and Curtis J. Humphreys, J. Research Nat. Bur. Stand. **38**, 211 (1947).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

An Automatic Recording Infra-Red Spectrophotometer

W. S. BAIRD, H. M. O'BRYAN,* GEORGE OGDEN, AND DOROTHY LEE
Baird Associates, Inc., Cambridge, Massachusetts

(Received July 18, 1947)

An automatic recording spectrophotometer for the infra-red is described, which employs a double-beam source, Littrow spectrometer, fast bolometer, and high speed pen- and drum-recording system. The sample and reference beams are alternately chopped at 10 cycles/sec. at the entrance slit to the spectrometer. The amplified a.c. signal on the bolometer is used to drive a variable comb diaphragm in the reference beam. The spectra are registered on a recorder by a pen mechanically linked to this comb.

The spectral region covered is between 2 microns and 16 microns with a sodium chloride prism. The range can be extended to about 26 microns by the use of a potassium bromide prism. The instrument is designed for high speed analytical work, producing permanently recorded spectrograms of gases, liquids, and solids on accurate charts at intervals of 18 minutes. Percent transmission is recorded with an accuracy of ±1 percent, and wave-length with an accuracy of ±0.025 micron. Records are reproducible to within ±1 percent transmission and ±0.01 micron.

Special features are its high recording speed, stability and freedom from drift with the "optical-null" principle, linearity and ease of operation.

A. INTRODUCTION

THE need for an instrument capable of recording the infra-red absorption spectra of organic compounds has been recognized for some years. The principal limitations on the development of such an instrument have been: (1) lack of a fast and sufficiently sensitive receiving element, and, (2) inadequacy of electro-mechanical recording systems. The two are necessarily interdependent since improvement in either helps to close the gap between performance and non-performance.

Progress in recent years in electronic amplifiers and servo-mechanisms on the one hand, and heat-detecting elements on the other, has been such that it is now possible to obtain an accurate record of an absorption curve in the analytically useful region of the infra-red spectrum in considerably less than one hour.

Most of the work on this problem has been concerned with providing a mechanical scanning arrangement for a conventional infra-red monochromator, amplifying the output as received on a thermocouple or bolometer, and producing a record on a standard recording instrument. Some excellent results have been obtained, but all instruments operating on these principles have two major inherent faults: (1) drifts caused by changes in ambient temperature, amplifier gain, etc., and (2) the fact that *all* changes in light intensity are recorded, whether caused by absorption by the sample, atmospheric absorption, or the change in intensity of source radiation as the spectrum is scanned.

The second of these deficiencies makes it necessary to run a reference spectrum without a sample immediately preceding or following the sample run, in order to determine the effect of the sample on the transmission curve of the instrument. The presence of ambient and electronic drifts increases the probability of error in comparing the two records. Accurate results can be obtained only by the exercise of extreme care in electronic design and construction, and elaborate precautions to insure stable ambient conditions in the vicinity of the sensitive element of the apparatus.

The development of high speed receiving elements has minimized the drift problem by permitting the use of interrupted radiation, the receiver output being amplified by an a.c. system sharply tuned to the interrupter frequency. This, however, does not remove the uncertainty in comparing the two records, assumed to be identical except for the absorption of the sample, taken at time intervals of one-half to one hour. This difficulty can be overcome only by comparison of the sample *versus* no-sample condition either simultaneously, or at such minute intervals of time that no change of conditions of sufficient magnitude to cause an error could have occurred. The obvious method of achieving this is to employ two beams of radiation, as nearly

* Now with the Joint Research and Development Planning Board, Washington 25, D.C.

identical as possible except that one is passed through the sample under investigation, and cause them to fall alternately on the receiving element at a frequency of several cycles per second.

The principal work on an optical, mechanical, and electrical system that provides nearly instantaneous comparison between the sample and reference beams was done by J. D. Hardy[1] and Norman Wright.[2]

The method used has an additional advantage in that it lends itself readily to shuttering one beam of radiation to compensate for absorption of the sample in the other beam. With a servomechanism arranged to drive continuously to a null condition, the system is independent of variations in source intensity, receiver sensitivity, and amplifier gain. The light passed by the comb-shaped diaphragm as described by Wright[2] and originally by Lehrer[3] is completely linear with respect to its angular position, so that a recorder pen mechanically coupled to this comb will plot percent transmission (or absorption) directly on a linear scale.

B. CONSTRUCTION AND OPERATING PRINCIPLE

It was decided to build an infra-red spectrophotometer on the "optical-null" principle which would give accurate spectra at as high a speed as possible without sacrificing resolution and reproducibility.

The instrument employs a double-beam source, 10-cycles/sec. alternating interrupter, Littrow spectrometer, platinum-strip bolometer, high gain, low noise voltage amplifier, synchronous rectifier, power amplifier, comb diaphragm, and a pen and drum recorder.

Double-Beam Source

The source proper is a $4'' \times 2'' \times \frac{1}{4}''$ Globar unit operated at 1400°K to give the energy peak at about 2.2 microns. The Globar is contained in a tubular water-cooled housing with two windows spaced about 120° apart. The entire housing, including cooling coils, is chromium plated to reduce radiation from the surface to a minimum.

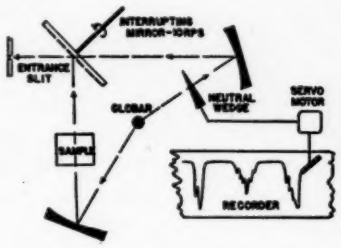

FIG. 1. Null system.

Radiation from the two windows is focused in the plane of the interrupter mirror by two spherical front-surface aluminized mirrors, with 10-cm diameters, and 20-cm focal lengths. Figure 1 shows the schematic arrangement.

Interrupter

The chopper is a mechanical interrupter with a circular sector wheel, half open and half plane mirror, located in front of, and at 45° to the entrance slit of the spectrometer. The sample beam is reflected by the mirror portion of the wheel, and the reference beam passes through the open portion. The sector interrupts each light beam at 10 cycles/sec.

Littrow Spectrometer

The spectrometer is of the Littrow design, a diagram of which is shown in Fig. 2. With this arrangement, the prism can be used twice, giving twice the dispersion of a single prism instrument.

Slits. The entrance and exit slits are driven by a cam and lever system which provides for bilateral opening of both slits simultaneously. The cam is designed to give approximately constant energy from the Globar over the 2- to 16-micron

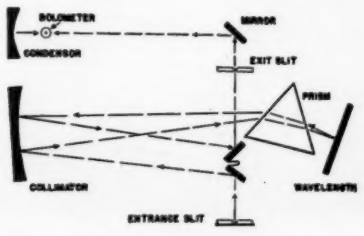

FIG. 2. Spectrometer.

[1] J. D. Hardy and A. I. Ryer, Phys. Rev. **55**, 1112 (1939).
[2] N. Wright and L. W. Herscher, "Double-beam, percent transmission recording infra-red spectrophotometer," J. Opt. Soc. Am. **37**, 211 (1947).
[3] E. Lehrer, Zeits. f. tech. Physik **23**, 169 (1942).

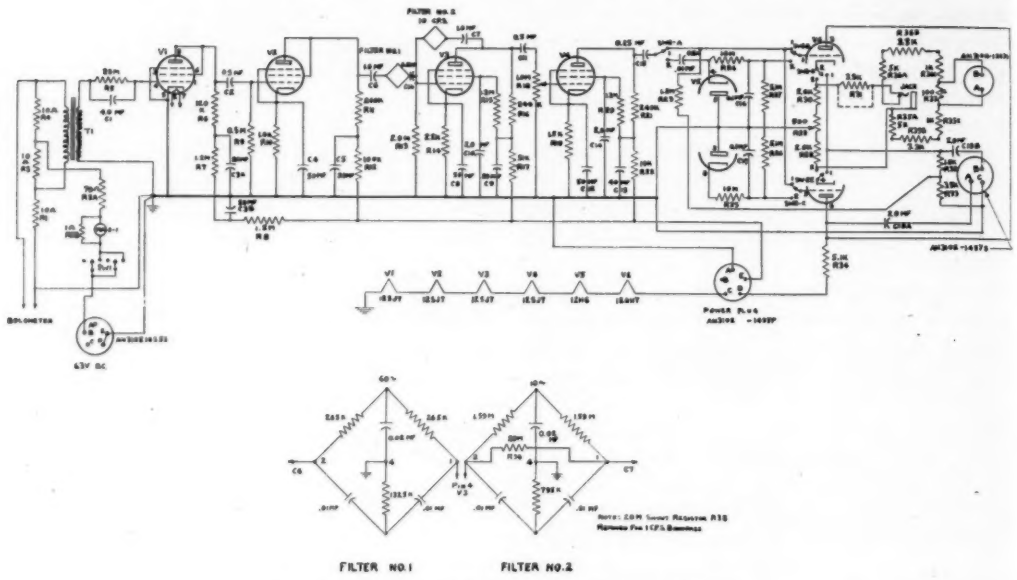

FIG. 3. Voltage amplifier—infra-red spectrophotometer.

range, this being the condition for most stable operation electronically. The slit openings vary from 0.002″ at 2 microns to a maximum of 0.090″ at 16 microns. The entrance slit is curved to compensate for the curvature of its image by the prism. Both slits are 38 mm high. The exit and entrance slit mirrors are two 45° plane, front-surface aluminized mirrors mounted on the same pedestal. The spacing between the mirrors is adjusted to give an angle of 2.3° between the entrance and exit beams.

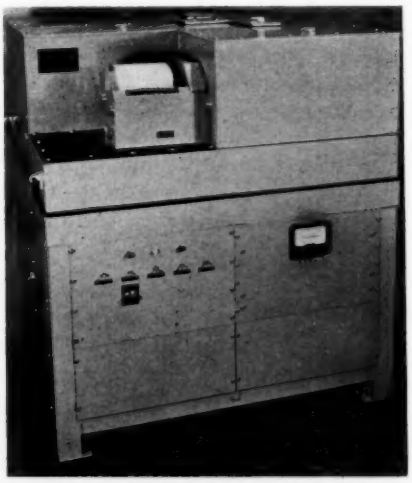

FIG. 4.

Collimator. The collimator is a 75-cm focal length, 13-cm diameter spherical concave mirror, front-surface aluminized, which is used 3½° (half-angle) off axis. This mirror is first used to send parallel light to the prism and then to focus the dispersed parallel light from the prism onto the exit slit. The system has an F ratio of about 10 ($f=75$ cm, and d, the diameter of the projected prism on the collimator, $=7.5$ cm).

Prism. The dispersing element is a 60° sodium chloride prism, 10-cm base, 8-cm height, kinematically mounted. The prism surfaces are flat to within 4 fringes of yellow light. The prism is adjusted for an angle of incidence of 52° to give minimum deviation at 10 microns.

The rocksalt prism can be replaced by a potassium bromide prism of the same size, and the range of the instrument extended to 26 microns. While the calculated optimum setting of the KBr prism would give minimum deviation at 18 microns, it was determined experimentally that no appreciable loss of resolution is suffered by placing it in exactly the same position as the

sodium chloride prism. Special slit and wave-length cams are required when prisms are changed, the cams being cut to give a linear wave-length scale and approximately constant energy from 12 to 26 microns.

To change prisms, it is necessary to remove the cover of the instrument, lift the rocksalt prism in its holder from the prism table, set the KBr prism, in its holder, on the table, remove the slit and wave-length cams and place the new cams. The slit and wave-length cams are screwed together and can be removed and replaced as a unit. The entire change-over requires about one hour.

Littrow Mirrow and Scan System. The Littrow or wave-length mirror is a front-surface aluminized plane mirror. The spectrum is scanned by the rotation of this mirror. The angle through which the Littrow mirror moves in order to keep the wave-length scale linear is determined by the wave-length cam and lever system.

Bolometer and Receiver System

The receiver system consists of a 45° plane mirror, and a front-surface aluminized spherical condenser of 9.9-cm focal length which focuses the exit-slit image on the bolometer. This is shown in Fig. 2.

The bolometer element itself consists of a thin platinum strip 7 mm long, 0.6 mm wide, and 0.05 micron thick, with a time constant of 0.05 second, a sensitivity of approximately 1 volt/watt at 10 cycles/sec., and a noise level of less than 5×10^{-9} volts. The bolometer strip is gold-blacked and has a resistance of 8 to 10 ohms at 50°C. The bolometer housing has a KBr window for transparency up to 26 microns. The housing must be evacuated to a pressure of about $\frac{1}{2}$ mm of mercury for most efficient operation. A more complete discussion of this fast, sensitive bolometer is being prepared by W. G. Langton of Baird Associates, Inc.

C. OPTICAL AND ELECTRICAL NULLING SYSTEM

The electrical system consists of a high gain low noise voltage amplifier, a mechanical rectifier, a power amplifier, the light-balancing mechanism, and the recorder.

The servo-cycle may be described as follows: In those regions of the spectrum in which the sample does not absorb infra-red radiation (that is, at 100 percent transmission), the sample and reference beams are of equal intensity and the resultant 10-cycles/sec. a.c. signal from the bolometer is zero.

At the characteristic absorption bands of the sample, the two beams will be of unequal intensity, and there will be a 10-cycles/sec. a.c. signal from the bolometer, the amplitude of which depends upon the difference in intensity of the two beams. This a.c. component is amplified about 50 million times by the voltage amplifier. (Figure 3.) Full-wave rectification is accomplished by two switches mounted on the interrupter shaft and synchronized with the interrupter mirror. The resulting d.c. signal, whose polarity is determined by whichever of the two beams is the more intense, is r-c filtered, converted into 60-cycle a.c. and amplified by

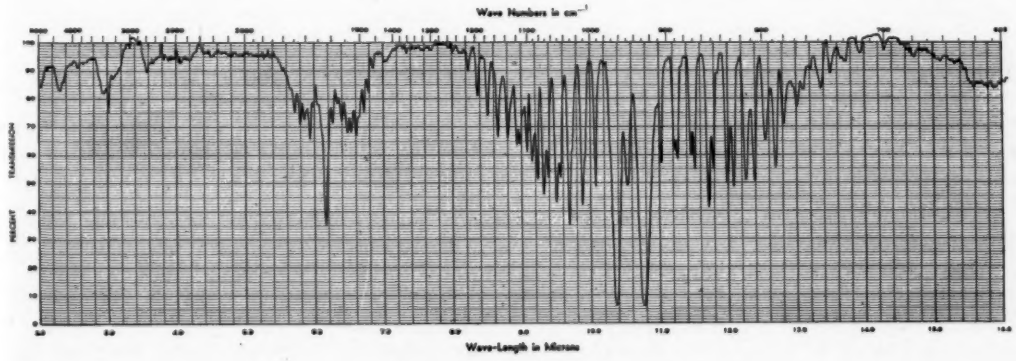

FIG. 5. Ammonia spectrum.

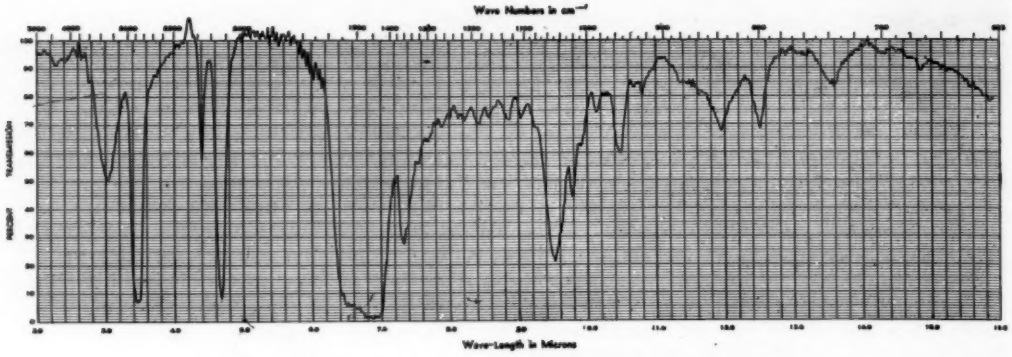

FIG. 6. Cholesterol in carbon disulfide; blank cell in reference beam.

the power amplifier (a Brown Electronik recorder amplifier).

The 60-cycle signal from the power amplifier is used to drive a Brown Electronik two-phase control motor. The phase and amplitude of this signal determines the direction and rate of motion of the compensating diaphragm or comb. Thus, when the amplitude is large, the comb travels at a rate proportional to the magnitude of the signal, giving a steeper band than when the amplitude of the signal is small, although in both cases the final percent absorption (or transmission) may be the same. The comb has six triangular teeth to provide a linear response to the change in intensity and continuously balances the system optically by trimming the reference beam, thus reducing the a.c. signal from the bolometer. When the two beams have been equalized, the resulting signal from the bolometer again has no alternating component and the system is optically and electrically at null.

The pen is mechanically linked to the comb and continuously records the position of the comb in percent transmission on the recorder chart. Twelve minutes are required to record 2 to 16 microns and six minutes to return to the starting position. The recording drum is rotated by the same motor which operates the slit and wave-length systems.

Figure 4 shows a complete instrument. The optical parts are all bolted to a sturdy welded frame which is stress-annealed after welding and cross-braced to eliminate distortion. The frame rests on four specially constructed shock mounts employing coil springs damped by soft rubber. These mounts are designed with a natural period of 3 cycles per second and greatly attenuate any frequency to which the instrument is sensitive.

The lower supporting part of the instrument is essentially a steel table. The voltage amplifier, electronic power supply, infra-red source power supply, power amplifier, battery for the bolometer supply, and control panel are all mounted in this section. The side panels of the supporting table are easily removable, providing ready access to the electrical components. The instrument is assembled and packed as a single unit, but the frame and table can be separated if necessary for moving through doorways.

The amplifier is suspended from springs with a natural frequency of 3 cycles/sec. to reduce microphonics due to shock and vibration.

A vacuum line for the bolometer is built into the instrument, with a connection at the rear for attaching the vacuum pump. Connections for the water supply to the source are provided in the same location.

The upper section of the instrument is tightly gasketed to permit a dry atmosphere to be maintained inside when an air conditioned room is not available for installation of the instrument.

The external controls have been kept to the minimum possible, being restricted to two switches, one push button, and one pilot light, all mounted on the front panel of the recorder assembly. The switches control the main power supply and the scanning motor, the pilot light indicates that the power is on, and the push

button starts the recorder and scanning mechanism, which stops automatically when a complete spectrum has been run and the starting position regained.

D. PERFORMANCE

The performance of this instrument can be seen from Fig. 5 which shows the spectrum of ammonia at 100-mm pressure with a 5-cm rocksalt cell. The spectrum reproduces better than .01 micron in wave-length and within ±1 percent in transmission.

The resolution to be expected from the present instrument depends on the setting of the prism, focal length of the collimator, slit widths, the recorder resolution, the energy available from a blackbody (Globar) and noise limitations.

With the prism set for a 52° angle of incidence, the dispersion was calculated as follows: Three points of a narrow wave-length band was examined $[\lambda-(d\lambda'/2)$, λ, and $\lambda+(d\lambda/2)]$. λ was first traced through the prism in the forward direction with Eqs. (1)–(3).

$$\sin(r)=\sin(i)/n, \tag{1}$$

$$(i')=A-(r), \tag{2}$$

$$\sin(r')=n\sin(i'), \tag{3}$$

using $(i)=52°$. The index (n) of rocksalt was obtained from Schaefer and Matossi.[4] λ was then traced in the same direction through the prism using $(i)=54.3°$. This 2.3° difference is the angle of separation between the exit and entrance beams. The Littrow mirror was then set perpendicular to the average of the two values of (r') thus calculated.

λ_1 or $\lambda-d\lambda/2$ was then traced through the prism once using $(i)=52°$ to find $(r\lambda_1')$. Then the angle of incidence $(i\lambda_1)$ of λ_1 on the prism after reflection by the Littrow mirror is given by

$$i\lambda_1=2r\lambda'-r\lambda_1', \tag{4}$$

where $(r\lambda')$ is the average (r') found above by tracing λ through the prism twice in the forward direction, and $(r\lambda_1')$ is the angle of emergence of λ_1 after being traced through the prism once in the forward direction. With $(i\lambda_1)$, trace λ_1 back through the prism to find a new value of $(r\lambda_1')$ to be compared with 54.3°. λ_2 or $\lambda+d\lambda/2$ is handled in the same way, where

$$i\lambda_2=2r\lambda'-r\lambda_2', \tag{5}$$

and a new value of $(r\lambda_2')$ found. The angular dispersion $dr'/d\lambda$ is simply $(r\lambda_1')-(r\lambda_2')/d\lambda$.

The linear dispersion in microns per millimeter, or $d\lambda/dy$, is defined as follows:

$$d\lambda/dy-d\lambda/\sigma dr', \tag{6}$$

where σ is 750 millimeters or the focal length of the collimator. This procedure gives the dispersion at one wave-length. To find the complete dispersion curve, several points in the 2- to 16-micron region must be calculated by the above method.

The optical resolution, $\Delta\lambda_D$ is defined as

$$\Delta\lambda_D=d\lambda/dy\times\text{the slit width}. \tag{7}$$

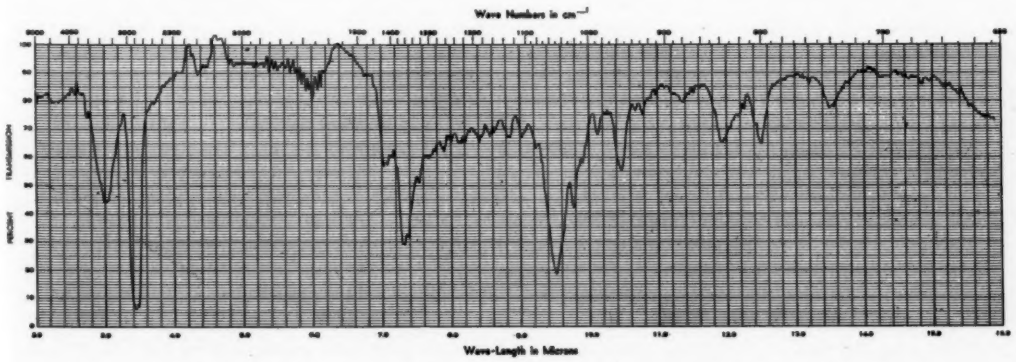

FIG. 7. Cholesterol in carbon disulfide; carbon disulfide in reference beam; compare with Fig. 6 and note cancellation of CS_2 bands.

[4] Schaefer and Matossi, *Das ultrarote Spectrum* (J. Springer Verlag, Berlin, 1930), p. 47.

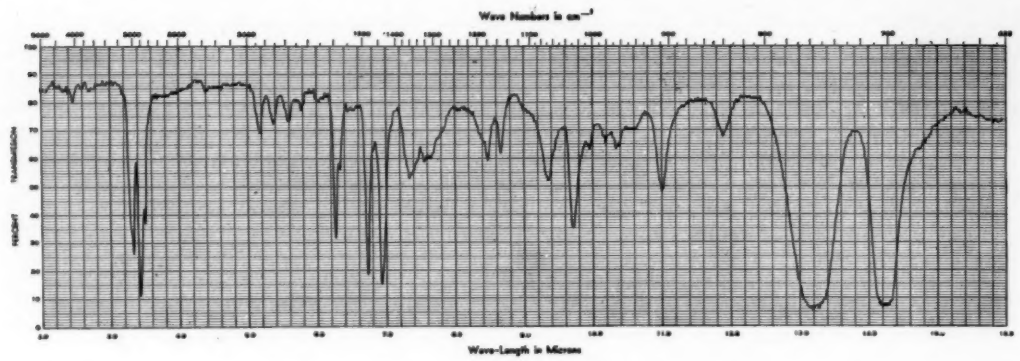

FIG. 8. Absorption spectrum of 0.025" polystyrene film.

With slit widths of 0.1016 mm, 0.457 mm, and 1.27 mm, $\Delta\lambda_D = 0.013\mu$ at 5μ, 0.032μ at 10μ, and 0.06μ at 13.5μ.

The optimum scanning speed was determined experimentally by successively scanning the ammonia spectrum at slower and slower speeds until no further gain in resolution could be obtained. The over-all resolving power of the instrument is therefore assumed to be limited by the optical resolution of the spectrometer rather than by the recorder mechanism. A two percent change in recorded transmission between two bands has been taken as evidence that the two bands are resolved. By this standard, Table I

TABLE I.

λ	$\Delta\lambda$ Calculated	$\Delta\lambda$ Experimental
5.0μ	0.013μ	0.03μ
10.0μ	0.032μ	0.04μ
13.5μ	0.053μ	0.06μ

gives a comparison between the calculated resolution and the experimental resolution at three points. The experimental resolution as shown in Fig. 5 is about $.035\mu$ at 5μ and better than $.05\mu$ at 10μ. For example the two bands at 9.47 and 9.53 microns are resolved by 4 percent.

Other performance examples can be seen in Figs. 6, 7, and 8. The double-beam system is particularly efficacious in the measurement of the absorption curve of a solid in solution, in that the solvent can be placed in the reference beam. The bolometer thus sees only the difference between the two beams, this difference being the spectrum of the solid. Figure 6 shows the combined spectrum of cholesterol in carbon disulfide, i.e., with a blank cell in the reference beam. Figure 7 shows the same mixture, but with carbon disulfide in the reference beam. The method can be applied to the measurement of one constituent in the presence of other substances, by placing all except the one to be measured in the reference beam.

Transparent solids such as polystyrene (Fig. 8) are handled by simply leaving the reference well empty, i.e., with air as the standard.

E. CONCLUSIONS

The Baird Associates, Inc., double-beam infrared spectrophotometer was designed as an analytical instrument. As such, in comparing this instrument and its performance with instruments which operate from point to point, or by deflection, numerous factors must be taken into consideration.

While the double-beam instrument provides a facility for doing many things that were not possible with the single-beam instruments, there is still some interference from water vapor and CO_2, and the user must be aware of the limitations which this imposes on some problems. The signal-to-noise ratio is still lower than is desirable and this problem is being worked on. Although the range can be extended with a potassium bromide prism, or expanded in the lower wavelength region with a lithium fluoride prism, these range changes are not usual and in general some time is necessary to check the wave-lengths and

slit settings after they have been made. Further, the instrument does not have the flexibility needed by some research workers, there being no simple way to do reflection problems, nor is there an easy way to do small solid samples. These deficiencies are a result of compromises in design necessary to give the resulting performance which has been shown with the double-beam standard instrument.

ACKNOWLEDGMENT

The authors wish to acknowledge the cooperation of the following: Dr. Norman Wright and Leonard Herscher of the Dow Chemical Company; Dr. J. D. Hardy of the Cornell Medical Center; practically the whole of the Baird Associates, Inc. organization, but in particular R. O'B. Carpenter, Wm. G. Langton, and R. V. Meyer.

Selective Infra-Red Gas Analyzers*,†

W. G. FASTIE** AND A. H. PFUND
Department of Physics, The Johns Hopkins University, Baltimore, Maryland

(Received June 11, 1947)

A summary of infra-red methods of gas analysis is presented. The experimental work which led to the development of an instrument for service applications is described. The properties of this final instrument are outlined, together with a résumé of performance characteristics for a number of applications.

FOR many years the infra-red absorption characteristics of gases have been used for quantitative analysis of gas mixtures. Despite the complexity of the optics and of the detecting system of an infra-red spectrometer, this method is widely used today for a large variety of laboratory and industrial gas-analysis problems. Because of the pronounced success of the spectrometer in handling gas problems, efforts have been made by a number of investigators to develop simpler infra-red equipment for such applications.[1–5]

In 1941, the government contracted with Johns Hopkins University to study methods of infra-red gas analysis and to develop instruments based on the most promising methods, the important requirements being high sensitivity, portability, and selectivity (ability of the instrument to detect one component in a mixture and ignore all others). The purpose of this paper is to first review and compare the various methods which were considered, and then to discuss the experimental work which led to an instrument which was used for a number of Service applications.[6]

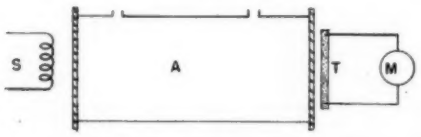

FIG. 1. Non-selective gas analyzer comprising a source, a sample space, and a radiation detector.

I. COMPARISON OF METHODS OF INFRA-RED GAS ANALYSIS

Non-Selective Instruments

In contrast to the infra-red spectrometer, which requires precision optics and mechanics, low optical aperture, and highly sensitive detecting systems, the gas analyzer shown in Fig. 1 is simple, rugged, and very sensitive. It consists of a source S, a sample space A, a detector T, and a relatively insensitive meter or recorder M. It is completely non-selective, any infra-red absorbing gas which enters the space A reduces the radiation reaching T and, therefore, produces a change in meter output.

Use has been found for such simple instruments in the study of the breakdown characteristics of gas adsorbing materials. Different types of gases can be studied without altering the instrument. Although its use is quite limited, it is worth noting that the non-selective instrument is as sensitive as any other method of infra-red analysis.

The simple instrument shown in Fig. 1 can be made partially selective by the use of special windows and special receiving areas on the detectors. For example, an instrument was constructed with a magnesium-oxide powder filter on the source window and with a gold-backed glass receiving area on the thermopile. This

* The work reported in this paper was done under OSRD Contract OEMsr 178, Section 17.1 NDRC. Final report OSRD No. 5674, October 31, 1945. It was reported at the Cleveland meeting of the Optical Society.

** Now in Research Department, Leeds and Northrup Company, Philadelphia, Pennsylvania.

† This article was submitted by the senior author without the knowledge of Professor Pfund that it was intended for a special issue of the Journal.

[1] H. Schmick, U. S. Patent No. 1,758,088.

[2] A. H. Pfund, Science **90**, 236 (1939).

[3] Pfund and Gemill, Bull. Johns Hopkins Hosp. **67**, 61 (1940).

[4] K. F. Luft, Zeits. f. Physik **24**, 97 (1943).

[5] N. Wright and L. W. Herscher, J. Opt. Soc. Am. **36**, 195 (1946).

[6] Instruments for use by the Armed Services were engineered and built by Leeds and Northrup Company under a separate OSRD Contract OEMsr 1035, Section 17.1 NDRC.

instrument was only sensitive to gases which absorbed in the 7–12μ region. However, solid materials have, in general, broader absorption bands than gases and are, therefore, not selective enough to be used as primary filters, but are most valuable for decreasing the effect of interfering gases in the highly selective instruments described below.

Selective Source

If the non-selective source shown in Fig. 1 is replaced with a hot jet of gas, as shown in Fig. 2, the resulting instrument is highly selective. The gas will emit in only the very narrow spectral region in which it absorbs, thus the detector will receive no radiation that a contaminant gas can absorb. For example, a hot jet of CO_2 emits no radiation which CO absorbs but emits strongly at approximately 4.3μ the wave-lengths at which cold CO_2 is strongly absorbent.

This method has three serious drawbacks. The instrumental problem of maintaining a stable hot jet of pure gas is not an attractive one. Many gases decompose at temperatures below an efficient operating temperature. The emission bands of a hot gas are broader than and somewhat spectrally shifted from those of the cold gas sample, resulting in a loss of selectivity and sensitivity.

Selective Detector

This is the method previously reported.[2,3] As shown schematically in Fig. 3, the source and sample cell are the same as in Fig. 1. The detector *D* is a gas cell which contains thermocouples for measuring the temperature difference between the gas in the cell and the walls of the cell. If it is desired to measure CO_2, the detector cell *D* is filled with pure CO_2. Radiation from the source heats this gas, but the only radiation effective is in the very narrow spectral region in which CO_2 absorbs. Therefore, CO or other absorbing gases in the sample cell which do not have absorption bands in common with CO_2 cannot remove any of the energy which heats the gas and cannot affect the temperature of the CO_2 in the detector cell. However, when CO_2 enters the sample cell it absorbs the very radiation which heats the gas in the detector cell and, therefore, causes a change in the output of the detector-cell thermocouples.

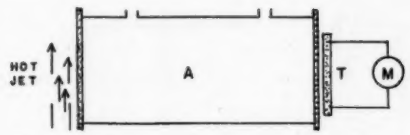

FIG. 2. Selective gas analyzer with source of hot gas, sample tube, and detector.

This method is the most selective one which has been studied. However, because of design difficulties a portable instrument based on this technique was not developed. An alternative method of detecting the amount of radiation absorbed by the detector cell is to measure the pressure in the cell. If an intense beam of intermittent radiation enters a cell containing CO_2 and having a flexible diaphragm in one side, the diaphragm will vibrate to such an extent that it can be detected with a stethoscope. By using this diaphragm as one plate of a condenser, an electrical signal can be obtained. K. F. Luft[4] has developed an industrial instrument using such detectors. These instruments were used for industrial-control purposes in Germany during the war.

Negative-Filtering Method

This method was patented by Schmick[1] in 1930. It was apparently ignored until recent interest in infra-red methods produced several rediscoverers. It was suggested to the writers by a classified British Scientific Office Report. It is the method upon which the Leeds and Northrup instruments were based. Baird Associates, Cambridge, Massachusetts, produce an infra-red gas analyzer based on this method, following the design of Wright and Herscher.[5]

As shown in Fig. 4, the source and sample tube of Fig. 1 are used, but two detectors are now used

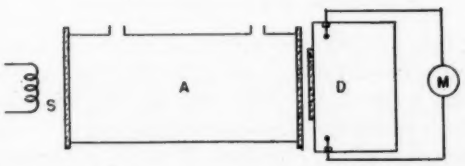

FIG. 3. Selective gas analyzer with source, sample tube, and selective detector.

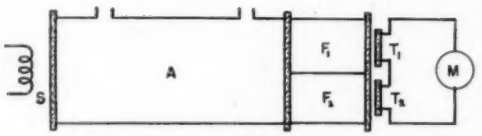

FIG. 4. Selective gas analyzer with source, sample tube, split negative filters, and differential detector.

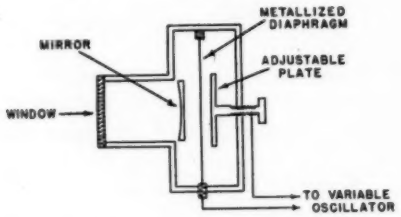

FIG. 5. Selective manometric detector.

and there is a gas-filter cell in front of each detector. The two detectors are connected in series opposition so that identical radiation signals on each detector produce no net thermopile output.

Let us assume that filter cell F_1 is filled with CO_2, and cell F_2 is filled with a non-absorbing gas. The detector T_1 does not receive any energy which CO_2 absorbs because the heavy filter of CO_2 in F_1 has removed this energy. The detector T_2 is sensitive to CO_2 radiation and responds to the entry of CO_2 in the sample cell. Both detectors are equally sensitive to other absorbing gases in the sample cell.

In short, this method consists of the use of two instruments, one which detects every absorbing gas including CO_2, the other which detects every absorbing gas except CO_2.

The most serious drawback to this method is the fact that two very large signals must be exactly cancelled. Success in applying this method depends upon the development of an instrument which can accomplish this accurate matching.

II. EXPERIMENTAL WORK

Selective Manometric Detector

This work started with an attempt to develop an instrument using a selective detector. At that time, however, there were no amplifiers available which could handle the low level output of the thermocouples. Therefore, detector cells using flexible diaphragms to measure the pressure in the cell were studied.

An experimental detector is shown in Fig. 5. The brass housing contains a rocksalt window through which the radiation passes, strikes the mirror, and is reflected out through the window again. The heated gas in the cell expands the metal-coated Pliofilm (0.003″) diaphragm, which is one plate of a condenser. With appropriate electrical circuits a change of capacitance corresponding to an extremely small radiation signal can be detected. When used with an infra-red source and 50-cm absorption tube, a concentration of 3-p.p.m. CO_2 could be detected. This signal was much larger than the electrical noise in the detector, the limit being set by mechanical and thermal disturbances.

At about this stage in the development, the General Motors d.c. chopper amplifier became available and the manometric detector was abandoned mainly because of the pressure of time. However, it may play a very important role in future gas-analysis developments.

Selective Thermoelectric Detector

Many designs of selective thermoelectric detectors were tried. The best experimental instrument based on this method is shown schematically in Fig. 6. The source was water-cooled, the radiation passed through the sample tube three times before entering the detector cell. The detector cell contained two rings of thermocouples. The thermocouple wires were lightly gold-plated to reduce the effect of scattered light. The two detectors were connected in series opposition so that the initial meter reading was zero. One side of the instrument was used as a compensator for voltage and temperature variations.

Because of the high optical efficiency of this instrument the aberrations caused a loss of

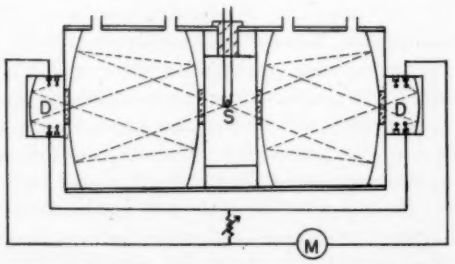

FIG. 6. Experimental instrument using matched double mirror sample tubes and opposed selective detectors.

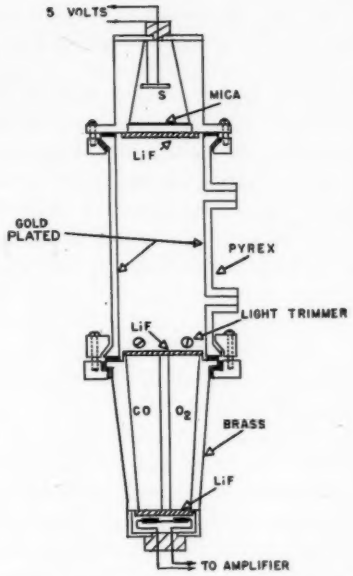

FIG. 7. Final experimental instrument using reflecting sample tube (5 in. long) and conical negative filters.

selectivity. A design which would combine the high fundamental selectivity of the method with the high sensitivity which was required for Service problems did not appear to be available.

Negative Filtering Instrument

At the time that the negative filtering method was suggested, all of the optical and electrical parts required to test the method were available. In a few days it was evident that an instrument based on the method was as sensitive, more selective, and more stable than the previous instruments which had been tested.

The first tests were made with instruments of the type shown in Fig. 6, with the detector cell replaced by a gas filter in front of a radiation thermopile. One filter was filled with CO_2 and the other with O_2.

The gas sample was simultaneously passed through both sample tubes. As explained above, one thermopile detected all absorbing gases and the other detected all absorbing gases except CO_2.

Although the double construction partially compensated for ambient temperature changes, temperature differentials across the instrument produced spurious indications, as did any local deterioration of window or mirror surfaces. These

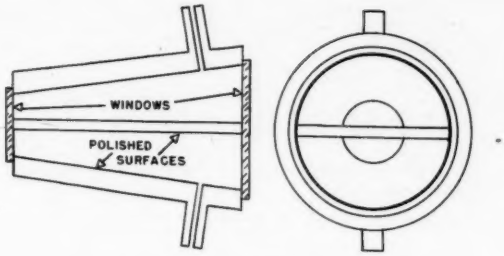

FIG. 8. Conical gas filters.

faults indicated the need for a single sample tube and proximity of the two detectors. The most difficult requirement which this design needed to meet was the high optical aperture of the double instrument (f:1.0).

Final Instrument

The instrument which finally evolved is shown in Fig. 7. The source is in a conical polished, air-cooled housing. The source is a flat coil of oxidized Nichrome operated at about 450°C, requiring 10 watts. The sample tube is of Pyrex, inside surface gold-plated, length 5 in., diameter $1\frac{1}{2}$ in.

The constructional details of the split-conical filter cells is shown in Fig. 8. These filter cones are made of glass or brass and their inside surfaces are polished and gold-plated. One side of the filter is filled with oxygen and the other side with 100 percent of the gas for which an analysis is desired. In front of each half of the cone is an adjustable screw which is used as a light trimmer.

The infra-red radiation leaving the source housing, reflects down the sample tube and is divided in two by the split cone, which has the dual action of a double filter cell and a radiation concentrator. The two high energy density beams leaving the cone, f all on the two halves of a

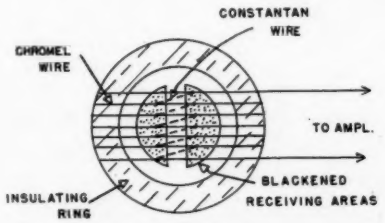

FIG. 9. Differential thermopile.

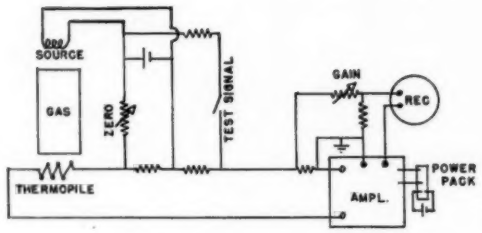

FIG. 10. Schematic diagram of electrical circuit for negative-filter gas analyzer.

differential thermopile which measures the difference in energy between the two beams. The design details of this differential thermopile are shown in Fig. 9.

There are many advantages in the use of the light-funnelling optical system just described. Because of the high infra-red reflectivity of gold the effective optical aperture is very high (f:0.7). This allows the source to be run at such a low temperature (450°C) that the major portion of the source radiation is in the important 3 to 10μ region of the spectrum. A large range of sample-tube lengths can be used. The simplicity of the components is an attractive production feature. The differential type of thermopile, surrounded by a heavy brass housing, is well compensated for temperature differentials. The sample tube has a small volume, and contamination spots on windows and tube walls produce equal signals in each radiation beam.

For the purposes of this discussion it is not necessary to outline all of the details of the amplifying and recording apparatus. The block diagram in Fig. 10 shows the main features. The output of the thermopile is fed into a d.c. chopper amplifier. The output of the amplifier is converted to d.c. with an output chopper in phase with the input chopper. Some of the output is inversely fed back to the input for gain stabilization. Variation of this feedback is used as a gain control. A potentiometer circuit is used for zero adjustment. The usual range gives a full-scale recorder deflection for 100 μv input signal.

Selectivity

Very careful matching of the two radiation beams is required in order to realize the full selectivity of the negative filtering instrument. This intensity matching can be adequately accomplished with the light-trimmer screws in front of the gas filters. Since one filter absorbs radiation, the two signals are not equal, resulting in an initial signal which is balanced out by a bucking potential in the input circuit of the amplifier. The initial signal must not be balanced out with the adjusting screws, since this would require reducing the radiation entering the oxygen cone non-selectively, i.e., equally at all wave-lengths. This would make the instrument slightly sensitive to all infra-red absorbing gases.

An instrument can be optically balanced before the filter cone is filled with absorbing gas, or afterwards, by testing with a contaminant gas and adjusting the light trimmers until no signal results from introduction of the contaminant into the sample cell.

Because the cones cannot be degassed, rem-

GAS DETECTED	RANGE	CONTAMINATING GASES	SENSITIVITY LIMIT	CAUSE OF LIMITATION
CO	0-1%	GUN FUMES	.005%	INSTRUMENTAL
CO_2	0-10%	PRODUCTS OF BURNING GASOLINE	.1%	CONTAMINANTS
CO	0-5%	PRODUCTS OF BURNING GASOLINE	.05%	CONTAMINANTS
HCN	0-1.5%	2% H_2O, 2% CO_2	.03%	CONTAMINANTS
CO_2	0-.1%	2% H_2O	.002%	INSTRUMENTAL
H_2O	0-3%	.04% CO_2	.01%	INSTRUMENTAL

FIG. 11. Résumé of performance of negative-filter instruments.

nants of CO_2 and H_2O in the cones can produce output signals if these gases are present as contaminants in the gas sample. As little as 20 p.p.m. of either of these gases in one of the cones causes measurable loss of selectivity. In order to minimize this effect, the filter cones and the gases with which they are filled must be uniformally treated so that equal traces of the gases are present in each filter.

Some problems involve detection of a gas in a mixture containing one or more contaminants with overlapping absorption bands, or with such high concentration of contaminants that the optical matching is inadequate. In such cases it is necessary to place another gas cell, permanently filled with the contaminant, in both radiation trains. Such a filter removes the radiation which the contaminant absorbs and desensitizes the instrument to the contaminant. This cell increases enormously the selectivity of the instrument. Even in cases in which two gases have almost identical bands, the contaminant filter produces a high degree of selectivity.

In many cases remnant sensitivities to contaminants can be minimized by judicious tampering with the light trimmers. For example, if gas *A* gives a slight up-scale deflection and gas *B* a slight down-scale deflection, and if these gases will be present in the sample in a definite ratio, the trimmers can be adjusted until the expected contaminants give no net signal. Furthermore, if a group of contaminant gases have similar bands partially overlapping the gas for which the instrument is sensitized, the trimmers can be adjusted until the effect of the group of gases is removed. However, the instrument will then be sensitive to other gases. The technique used to obtain optimum selectivity depends upon the particular problem.

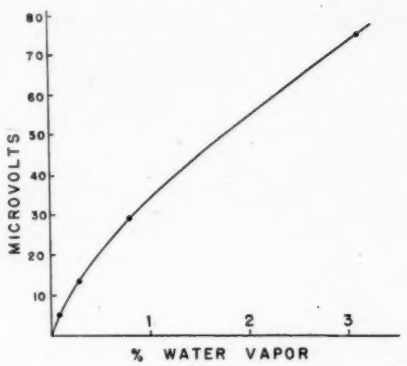

FIG. 13. Calibration curve for water vapor with 5-in. sample tube.

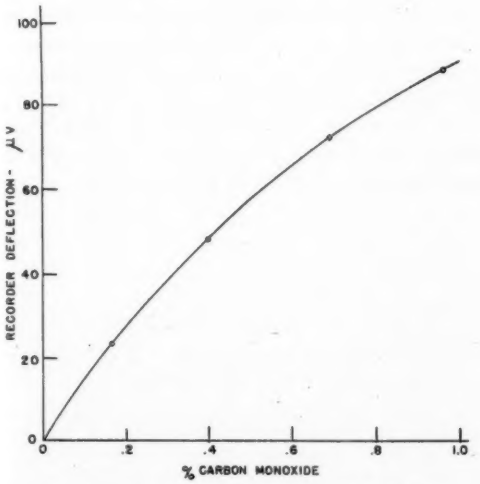

FIG. 12. Calibration curve for carbon monoxide with 5-in. sample tube.

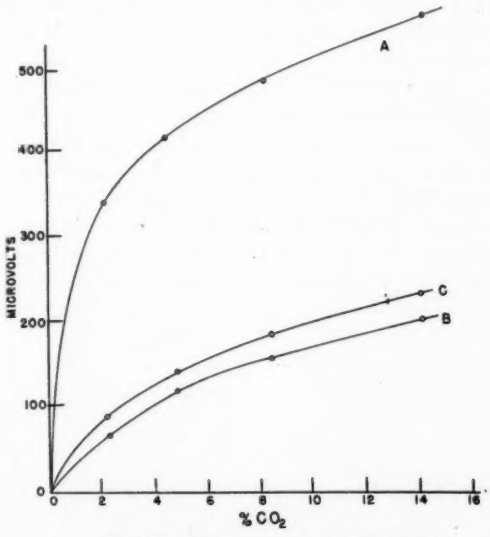

FIG. 14. Calibration curves for carbon dioxide.
A—5-in. sample tube.
B—5-in. sample tube with 5 percent CO_2 in oxygen side of cone.
C—$\frac{3}{4}$-in. sample tube.

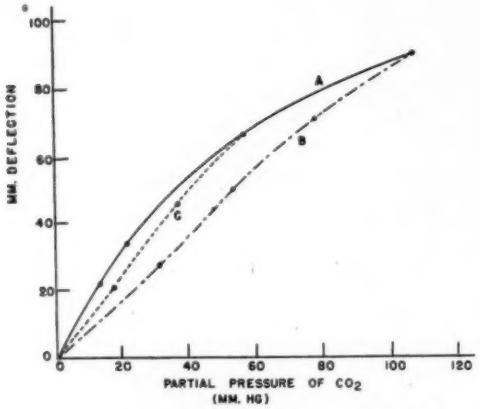

FIG. 15. Calibration curves for carbon dioxide.
A—total pressure constant.
B,C—percent CO_2 constant, total pressure varying.

Sensitivity

The limit of sensitivity of the instruments which have been built is determined by source-voltage fluctuation and by ambient-temperature fluctuations. Most Service applications have required battery operation and have not required long-term zero stability. Figure 11 is a résumé of the performance of the instruments under a large variety of operating conditions and gives a fair picture of the sensitivity and selectivity of the method.

Figures 12 and 13 show the calibration curve for CO and water vapor for a five-inch absorption tube. Curves *A* and *C*, Fig. 14, show the calibrations for CO_2 for 5-inch and $\frac{3}{4}$-inch sample tubes, respectively. Curve *B* of this figure shows the calibration curve obtained when one side of the filter cone is filled with pure CO_2 and the other side with 5 percent CO_2, 95 percent O_2. This produces a desensitized instrument with more nearly linear characteristics.

Figure 15 demonstrates the effect of pressure on the shape of the calibration curve, and is an example of the failure of Beer's law. This same effect can occur without change of pressure but with change in the kind of gas mixed with the CO_2. Therefore, it is necessary in some extreme cases to know pressure and approximate composition in order to determine the percentage of CO_2 or of another strongly absorbing gas. It is also necessary to avoid mixing gases in the filter cones. An instrument with one side of the cone filled with 33 percent CO_2, 67 percent O_2, and with the other side of the cone filled with 33 percent CO_2 and 67 percent CO, is not desensitized to CO_2 as might be expected, but is more sensitive to CO_2 than it would be if no CO_2 were in either side of the cone. This further manifestation of the effect of association on infra-red absorption bands forces the use of a separate contaminant cell for desensitizing the instrument to strongly absorbing high percentage constituents of gas mixtures.

ACKNOWLEDGMENTS

The work reported here extended over a four-year period and required the help of a great many people both in the development of the instrument and in its application. Messrs. C. Wilbur Pêters and Louis F. Drummeter of this laboratory have assisted in the program. All of the experimental apparatus was built by Mr. Calvin Krause in the Physics Department Shop. Messrs. J. R. Stewart and Frank B. Gray, Engineering Department, Leeds and Northrup Company, were responsible for developing the instruments which were supplied to various Service laboratories during the war.

An Estimate of Transparency of the Atmospheric Window 16 Mu to 24 Mu

ARTHUR ADEL
McMath-Hulbert Observatory, University of Michigan, Ann Arbor, Michigan
(Received May 6, 1947)

The effective solar radiation intensity, as observed through the atmospheric window between 16 mu and 24 mu, is determined by comparison with the observed radiation intensity of the earth's surface (thermocouple) below 13 mu. The ratio of the effective value to the value calculated for $T=6000°K$ is the transparency of the atmosphere to solar radiation within the interval 16 mu–24 mu. Opacity in this spectral region is due largely to water vapor and carbon dioxide. Provisional estimates of percent transmission, probably lower limits, are found to be:

Date	Precipitable water vapor	Spectral interval	Percent transmission
30 Sept. '41	1 cm	16 mu–19 mu	3.3%
24 Nov. '41	1 mm	16 mu–19 mu	14.3%
24 Nov. '41	1 mm	16 mu–22 mu	12.6%

The net loss of energy by the earth's surface, through the 16-mu–24-mu window, is small, but probably not negligible from a meteorological point of view.

THE infra-red solar spectrum has recently been extended from 14 mu to 24 mu through a new atmospheric window.[1] Systematic measurements of atmospheric transmission in this spectral region have not yet been made. It is, nevertheless, possible to provide an estimate of transparency of the window, and it is the purpose of this paper to do so. To this end, a comparison will be made between the emission spectra of a radiation thermocouple and the sun. These spectra have been observed in a manner which renders their areas (under the envelopes) in the same ratio as the effective radiation intensities of the two sources. The transparency of the new window to solar radiation will be given by the ratio of the effective radiation insity of the sun to the value calculated for $T=6000°K$.

Reproductions of automatically recorded solar spectra, 16 mu to 24 mu, have been published.[1] Present needs are adequately met by tracings of two of these, Figs. 1 and 2. The great band of carbon dioxide at 15 mu separates this window from the major region of transparency in the infra-red, which is bounded at its other extremity by the great band of water vapor at 6.3 mu.[2] Absorption by water vapor and carbon dioxide dominates the region from 16 mu to 24 mu; the influence of water vapor is strikingly evident in Figs. 1 and 2.

Reproductions of automatically recorded terrestrial emission spectra (net radiation spectra of thermocouple) are given in a paper on the infra-red emission spectra of the moon, earth, and atmosphere.[3] Again, it suffices for present needs to employ a tracing, Fig. 3. Although Fig. 3 depicts the KBr prismatic spectrum of net radiation from the blackened receiver of the thermocouple, it is by no means the complete curve of a *BB* at 300°K. Because of return radiation to the thermocouple from the water vapor, carbon dioxide, ozone, and nitrous oxide of the atmosphere, the curve of Fig. 3 is seriously deficient at 6.3 mu and beyond 13.1 mu, and departs somewhat from *BB* emission even between 11 mu and 13.1 mu.

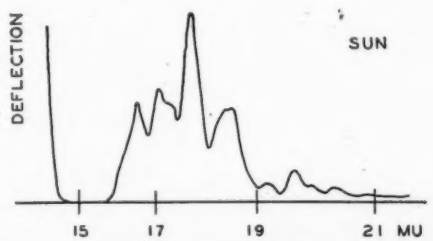

FIG. 1. KBr solar spectrum observed at Flagstaff, Arizona, September 30, 1941, through one centimeter of precipitable water vapor.

[1] Arthur Adel, Ap. J. **96**, 239 (1942). For a brief history of the problem see: (*a*) G. C. Southworth, J. Frank. Inst. **239**, 285 (1945); (*b*) *Annals of the Astrophysical Observatory of the Smithsonian Institution;* (*c*) J. Strong, J. Frank. Inst. **232**, 1 (1941).

[2] Arthur Adel and C. O. Lampland, Ap. J. **91**, 1 (1940); 481 (1940).

[3] Arthur Adel, Ap. J. **103**, 19 (1946).

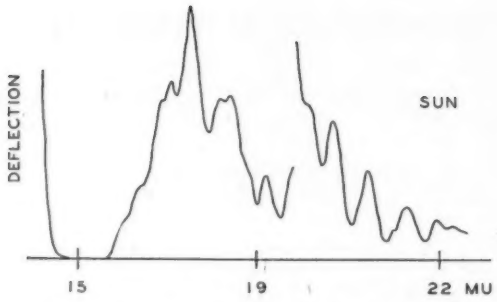

FIG. 2. KBr solar spectrum observed at Flagstaff, Arizona, November 24, 1941, through one millimeter of precipitable water vapor. Note scale change near 20 mu.

Following the minor correction necessitated by the departure from *BB* emission between 11 mu and 13 mu, and reconstruction of the envelope by bridging the absorption bands, the area under the curve, to 13.1 mu, should correspond to 47 percent of the radiation intensity of a *BB* at 300°K; that is, the area should equal 47 percent of that defined by the complete *BB* curve at 300°K. The solar and the thermocouple curves were secured with the same spectrometer and coelostat, the solar image covering the entire slit in the one case, the thermocouple radiating through the entire slit in the other, with the same coelostat aperture stop effective in both instances. That is to say, the thermocouple received energy from the sun, and, in turn, radiated its own energy, through the *same solid angle*. Hence, following correction for slit widths and amplification factors involved in the two experiments, the areas under the curves will be in the same ratio as the corresponding effective radiation intensities of the sources: sun and thermocouple. In units of the area under the thermocouple curve, corrected for atmospheric radiation, short wave-length end to 13.1 mu, the reduced values, of Table I, columns 1, 2, and 3, are obtained for the solar curves.

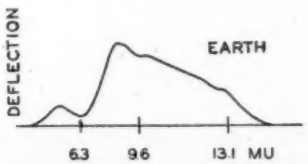

FIG. 3. KBr spectrum of thermocouple (net emission) observed at Flagstaff, Arizona, November 3, 1941.

The *BB* radiation intensity at 300°K is $J_{(0-\infty)300°K}=4.56\times10^5$ ergs/cm²/sec. And 47 percent of this amount lies below 13.1 mu; that is, $J_{(0-13.1\ mu)300°K}=2.14\times10^5$ ergs/cm²/sec. Application of this factor to the reduced, normalized areas of Table I yields the *effective* values, Table I, column 4, for the corresponding radiation intensities of the solar surface.

If we assume that the sun radiates as a *BB* in this spectral region, the calculated radiation intensities are:

$$J_{(16\ mu-19\ mu)6000°K}=4.80\times10^6\ \text{ergs/cm}^2\text{/sec},$$
$$J_{(16\ mu-22\ mu)6000°K}=7.37\times10^6\ \text{ergs/cm}^2\text{/sec}.$$

The ratio of effective radiation intensity to calculated radiation intensity is the percentage transmission of *solar radiation* through the new window, Table I, column 5.

Admittedly, these values are approximate, but it is probable that they are of the correct order. Furthermore, revision is likely to increase them, since thermocouple receiver blacks are seldom as efficient in the region of the new window as they are in the region of terrestrial emission.

It is important to consider the possible loss of energy from the earth's surface through the 16-mu to 24-mu window. About 17.5 percent of *BB* radiation, at 300°K, lies between 16 mu and 22 mu. Hence, the net escape of radiation from the earth's surface, while small, is perhaps not negligible, particularly when the air is both dry and appreciably colder than the surface.

The moon attains a temperature in the neighborhood of 100°C at the sub-solar point, and thus reaches a total radiation intensity more than twice that of the earth. Thirteen percent of *BB* radiation, at 373°K, lies between 16 mu and 22 mu. At times of dry atmosphere it should be possible to detect dispersed radiation arriving from the moon, as well as that escaping from the earth's surface, through the window 16 mu to 24 mu.

TABLE I.

Date	Spectral region	Area	Radiation intensity $J(\lambda_1-\lambda_2)$sun	Percent transmission
30 Sept. '41	16 mu–19 mu	0.75	1.60×10^6 ergs/cm²/sec.	3.3%
24 Nov. '41	16 mu–19 mu	3.20	6.85×10^6 ergs/cm²/sec.	14.3%
24 Nov. '41	16 mu–22 mu	4.35	9.30×10^6 ergs/cm²/sec.	12.6%

The temperature of the outer layers of Jupiter's atmosphere is approximately −110°C. The emission maximum of a *BB* at 163°K falls within the interval 16 mu to 24 mu, which also contains 21.7 percent of the total radiation intensity, but the radiation intensity of a *BB* at 163°K between 16 mu and 24 mu is merely 1/840 the corresponding value for the sun, and but little more than one-tenth the corresponding value for the earth. Nevertheless, the *integrated* Jovian radiation, 16 mu to 24 mu, might be detectable with large telescopes.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

The Diffuse Spectral Reflectance of Paints in the Near Infra-Red

J. A. SANDERSON
U. S. Naval Research Laboratory, Washington, D. C.
(Received July 16, 1947)

A perforated hemispherical reflector made of half of a Pyrex flask has been employed to collect diffusely reflected light from a surface to be measured on a photo-cell or thermopile placed at a focal point of the hemisphere conjugate to the position of the sample, thus providing adequate amounts of energy on the thermopile to permit spectral-reflectance measurements to longer wave-lengths than with usual reflectance spectrophotometers. Relative diffuse spectral-reflectance curves extending to 2.5μ are given for a number of "infra-red reflecting" paints, showing qualitatively the superiority of high index of refraction, large particle size, and low extinction coefficient in producing layers of pigments having high scattering coefficients in the near infra-red.

The use of the hemispherical integrator with a tungsten source and a Wratten 87 filter and with either a thermopile or photo-cell receiver is illustrated by measurements which show that the thermopile-Wratten 87 system, embracing wave-lengths 0.76 to 2.6μ, produced reflectance factors which more nearly approximated integrated reflection factors for sunlight between 0.76 and 2.4μ and more truly indicated the effect of variations in physical characteristics of the reflecting material than did photo-cell measurements embracing wave-lengths 0.76 to 1.2μ.

The use of a caesium oxide cathode, infra-red image tube in making rapid inspections of surfaces in the region of wave-length shorter than 1.2μ is described.

INTRODUCTION

THIS paper describes a simple reflectometer which was used in measuring the diffuse spectral reflectance and the integrated reflectance in chosen bands of the spectrum of a number of so-called "infra-red reflecting" paints and other surfaces in connection with a problem concerned with means of reducing the amount of sunlight absorbed by certain surfaces and the amount of radiation emitted by other, heated surfaces. It was realized that the notebooks of Professor A. H. Pfund of the Johns Hopkins University were replete with data pertinent to the achievement of special selective optical phenomena by the proper combinations of scattering particles with organic vehicles. He was asked, and he agreed, to serve as consultant to the Naval Research Laboratory on this problem. The work to be described in the present paper was aided considerably by his suggestions.

The term "infra-red reflecting" paint has come to be applied to any combination of pigments and vehicle with which surfaces may be produced which are of low reflectance in the visible spectrum and of higher reflectance in the near infra-red spectrum between about 0.7 and 2.5μ. Such surfaces are useful in cases where, for practical or aesthetic reasons, colored surfaces of low visual reflectance are required while at the same time minimum heating by absorbed sunlight is required. The long wave-length limit of interest in the present paper has been set at 2.5μ for a twofold reason: (a) the greater part of solar radiation reaching the surface of the earth lies at wave-lengths shorter than 2.5μ, and (b) at wave-lengths longer than 2.5μ nearly all films made up of irregularly shaped, granular pigments dispersed in organic vehicles are poor reflectors, regardless of the identity of the particles, because of the strong infra-red absorption by the

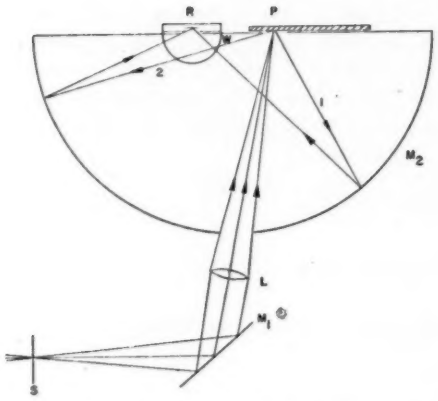

FIG. 1. Hemispherical integrator.

vehicle due to the long optical paths in it produced by multiple scattering of the incident light. (Aluminum paint is an exception, because the reflection from the "leafed" aluminum flakes lying underneath a thin layer of vehicle is primarily specular. Professor Pfund and the DuPont Company, working in behalf of the Navy, have achieved values of reflectance as high as 85 percent in the 8 to 13μ region of the spectrum in carefully formulated aluminum paints.)

The physical properties which layers of particles must possess if they are to scatter incident infra-red light, and not absorb it or transmit it, have been discussed by Pfund and Silverman,[1] Pfund,[2,3] and Barnes and Bonner.[4] Papers by Barnes,[5] Bryant,[6] and the author[7] described the selective infra-red reflection by layers of highly compressed powders.

The results of these papers showed that if a material which in its massive polished state is transparent to infra-red light is ground into small particles, and if these particles are collected near to each other in a loosely packed layer, they will scatter wave-lengths smaller than the particle size and transmit as collimated light wave-lengths larger than the particle size. (The effect is the basis of the well-known Pfund powder filters.) But if the material has high extinction coefficients in any part of the spectrum, the powder becomes quite black in that region. Thus, quartz, which shows metallic reflection at 9μ, arising from high extinction coefficients, becomes a strong absorber at 9μ when it is ground to a powder, and gold or aluminum when evaporated at a few millimeters pressure to form layers of small particles may be strong absorbers over a rather wide band of the spectrum. It is of interest that thin films of gold or aluminum black a few hundred angstrom units thick and made up of gold on aluminum crystals no larger than 200 to 300A in the biggest dimension, may be nearly complete absorbers of 15μ radiation. Pfund and Silverman[1] have pointed out that a simple light-trapping explanation is not sufficient to describe the behavior of small particles of high extinction coefficient. Furthermore, it has often been observed in this Laboratory that gold or aluminum blacks which absorb well in the visible spectrum may transmit well, say 90 percent, at 10μ, but the reasons for this behavior have not been established. It is also known that if gold black is heated so that it "sinters" it becomes shiny and reflects light, and that if aggregates of small particles such as $CaCO_3$ which absorb strongly at 6.56μ are compressed into a tablet presenting a smooth surface, they reflect light like a massive polished crystal of the same substance.[5–7] Hence, the scattering of light by an aggregate of small particles is related both to the roughness of the surface layer of particles and to the extinction coefficient.

It is readily apparent from these general properties of pigments that an infra-red reflecting pigment should be ground from material having high extinction coefficients for visible light and low extinction coefficients in the infra-red. Furthermore, the pigment particles should be large enough or the layer of them should be thick enough that a portion of the incident infra-red light will be scattered back and not be transmitted through the film without scattering.

The measurements of diffuse reflectance of a number of surfaces which illustrate some of the scattering properties of pigments will be described below.

[1] A. H. Pfund and S. Silverman, Phys. Rev. **39**, 64 (1932).
[2] A. H. Pfund, J. Opt. Am. **23**, 270 (1933).
[3] A. H. Pfund **23**, 375 (1933).
[4] R. Bowling Barnes and Lyman G. Bonner, Phys. Rev. **49**, 732 (1936).
[5] R. Bowling Barnes, Phys. Rev. **43**, 31 (1933).
[6] C. W. Bryant, J. Opt. Soc. Am. **24**, 139 (1934).
[7] J. A. Sanderson, J. Opt. Soc. Am. **30**, 566 (1940).

REFLECTOMETER

Conventional integrating spheres are not suitable for use in the infra-red because of the difficulty in producing a neutral, highly reflecting, diffusing inner surface for the sphere. The brilliant, white painted surfaces of photometric spheres suitable for use between 4000 and 7000A may be quite dark in the infra-red. Such a sphere coated on the inside with finely divided rocksalt or other infra-red-transparent material would be neutral if impermanent, but the intensity of light emerging from the sphere would be difficult to measure with thermocouples or bolometers.

A method of collecting virtually all of the light diffusely reflected by diffusing surfaces first employed by Paschen[8] and later by Royds[9] and by Coblentz[10] was used with success. A hemispherical reflector, Fig. 1, was made by sawing in half a Pyrex flask 16 cm in diameter. A hole was cut in the bottom of the reflector to admit light, and the inner surface was coated with aluminum by evaporation. A thin, blackened metal cover was provided containing apertures at P and R at equal distances from the center of curvature of the reflector and as near together as practicable. The sample to be measured and the reference sample were placed in turn over P; the intensity-measuring device, either photo-voltaic cell or thermocouple, was placed over R.

The reflector was mounted on a cork ring on a ring stand so that the top cover was horizontal, permitting exchange of samples without manipulation of any clamps or fasteners, the samples merely being laid in place. Light from the slit of the monochromator S was introduced into the reflector by way of the aluminized plane mirror M and the lens L which formed an image of the slit on the sample at P. Light diffusely reflected at P along any directions such as 1 and 2 within the hemisphere was brought to a focus at R since P and R were conjugate focii of the hemisphere. The image formed at R was not of high quality because of aberration, and the imperfect surface of the Pyrex flask, and the reflectance of surfaces differing widely in roughness, such as ground and polished glass, could not be compared. The instrument was used only in comparing painted surfaces with other similar surfaces which produced a similar distribution of scattered light over the hemisphere.

In the wave-length interval 0.4 to 0.8μ the intensity in the image at R was measured with a Lange photo-voltaic cell which was removed from its case so that the sensitive surface could be laid flat on the reflector cover at R approximately in the plane of the image and, hence, so that it was illuminated by light coming from all areas of the reflector. The photo-cell was connected to a suitable high resistance galvanometer.

For measurements at wave-lengths longer than 0.8μ the photo-voltaic cell was replaced with a four-junction copper-constantan thermopile. The blackened gold-foil receiving areas of this thermopile were each 4 mm square, producing a total area 4 mm$\times$16 mm which intercepted nearly all of the light in the imperfect image at R. The thermopile was also in the plane of the image so that it, too, was illuminated by light from all areas of the reflector. A thin hemispherical glass window W, made by sucking in a bubble at the end of a glass tube and subsequently grinding away the folded end of the tube so that the bubble could be removed, was placed over the aperture at P to exclude air currents from the thermopile and also to exclude long wave-length infra-red radiation emitted by the sample P because of heating by absorption of part of the incident light.

The hemispherical integrator was used in con-

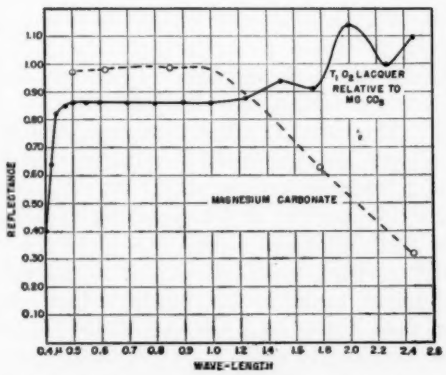

FIG. 2. Diffuse spectral reflectance of the TiO_2 lacquer reference standard relative to magnesium carbonate.
Note: The reference box in the upper right-hand corner of the graph should read TiO_2 lacquer relative to $MgCO_3$.

[8] Paschen, Ber. Berl. Akad. d. Wiss. **27** (April, 1899).
[9] T. Royds, Physik. Zeits. **11**, 316 (1910).
[10] W. W. Coblentz, Nat. Bur. Stand. Bull. **9**, 283 (1913).

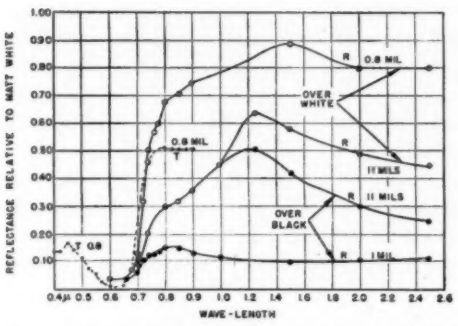

FIG. 3. Spectral reflectance of a pigment containing an organic toner.

junction with a Gaertner f:5.0 quartz monochromator employing a tungsten projection lamp as source and with the relatively insensitive thermopile connected to a Leeds and Northrup Type HS galvanometer. The system could be used with effective slit widths of about 0.1μ at 1μ and 0.15μ at 2μ.

REFERENCE STANDARD

The problem at hand was quite practical in that certain commercial infra-red reflecting paints were to be evaluated. Hence, it was decided to measure all of the paints relative to an arbitrary standard of good reflectance in the near infra-red: a metal plate was coated with a brilliant matte white lacquer containing titanium dioxide as its principal pigment, and all other surfaces were measured relative to this panel. The panel was "standardized" in a way by measuring its reflectance relative to magnesium carbonate. The resulting curve is shown in Fig. 2 together with existing data on the true diffuse reflectance of magnesium carbonate in the infra-red.[11] The point on the magnesium carbonate curve at 2.5μ was obtained from the smooth curve drawn through the 1.78μ point from (11) and the value 10.8 percent at 4.4μ given by Coblentz.[10] The irregular march of the TiO_2 curve at wave-lengths longer than 1.4μ is due to unresolved absorption bands in the organic vehicel. Since the same bands were present in the infra-red reflecting paints which were measured, they do not appear in reflectance measurements relative to the white lacquer.

[11] *Smithsonian Physical Tables*, eighth revised edition (1934), Table 434.

REFLECTANCE MEASUREMENTS

The results which have been chosen for presentation are those which illustrate in the smallest number of figures the wise variations in infra-red reflectance with physical nature of the pigment. Detailed information relating to the pigments cannot by given since the paints were commercial products and their compositions were not disclosed in detail. Each paint investigated was applied over both white and black undercoats on steel panels in order that the hiding power might be more readily evaluated.

Effect of Index of Refraction

Figure 3 shows the spectral diffuse reflectance of a paint containing as pigment an organic "toner" which possessed the poorest hiding power of any encountered in the investigation, the reflectance of a dry film thickness of 0.8 mil over white undercoat rising to 90 percent at 1.5μ while that of 1 mil over black undercoat reached a peak value of only 15 percent at 0.8μ and decreased to about 10 percent at longer wave-lengths. Thick, 11-mil films of this paint were prepared, but, as the reflectance curves show, even this thickness did not produce complete hiding of the underlying surface throughout the spectrum, the two samples being alike to 1.0μ but the one over white exhibiting higher reflectance at wave-lengths longer than 1.0μ.

Figure 4 shows the spectral-reflectance curves of a paint containing stibnite (antimony trisulfide) in dry film thickness of 2.3 mils over

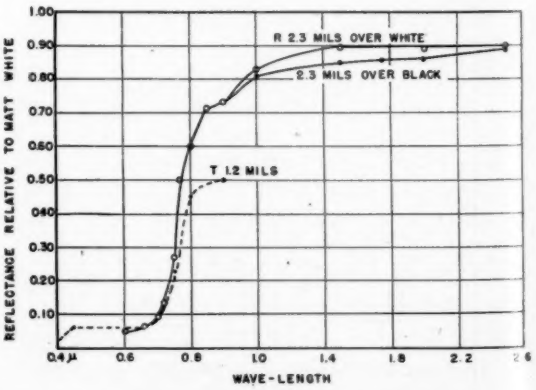

FIG. 4. Spectral reflectance of a stibnite pigment.

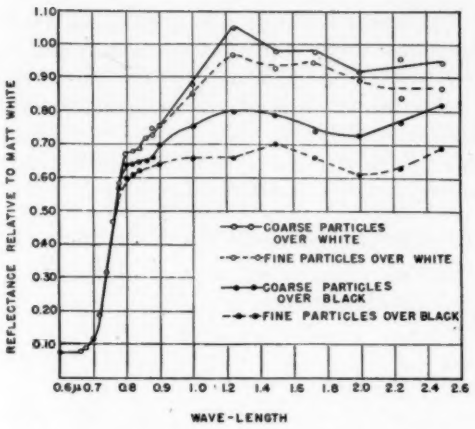

FIG. 5. Spectral reflectance of fine and coarse particles.

white and black undercoats. This paint possessed excellent hiding power, as the reflectance curves show, for at wave-lengths above 0.9μ the two reflectance curves differ by no more than five percent at any point, whether the scattering surfaces were over white or over black undercoats.

Pigments which behave in the manner shown by Fig. 3 are sometimes referred to as "infra-red transparent" while those for which Fig. 4 is characteristic are sometimes called "truly infra-red reflecting," the latter term implying a selective and special action of the pigment on incident infra-red light. These terms are misnomers, for the two types of pigments behave optically in the same way: the pigment particles act as scatterers of those wave-lengths to which the parent massive, polished material from which the pigment was ground is transparent, while they do not scatter wave-lengths to which the bulk material is opaque. Thus, the condition that an "infra-red reflecting" pigment must be ground of material which possesses high absorption coefficients in the visible spectrum and low absorption coefficients in the near infra-red is met by both the organic pigment and the antimony sulfide. But, the index of refraction of antimony sulfide being high, in the neighborhood of 4, while that of the toner unquestionably was lower and more nearly equal to the refractive index of the vehicle, the antimony trisulfide pigment possessed a higher scattering coefficient and, therefore, exhibited higher infra-red reflectance.

In Figs. 3 and 4 the dotted curves marked "*T*" are the spectral diffuse transmittance curves of specified thicknesses of the two kinds of pigments. These curves were obtained by allowing dispersed light from the exit slit of the monochromator to fall on a film of the material on a glass plate placed in contact with the sensitive surface of the Lange photo-voltaic cell. Thus, nearly all of the forward-scattered, diffusely transmitted light reached the sensitive area of the photo-cell. These curves are introduced to show their similarity to the reflectance curves, illustrating the identical role of the two types of pigments as scatterers of infra-red light rather than as selective reflectors or selective transmitters.

Effect of Particle Size

Figure 5 shows the spectral reflectance of two paints of identical (but unknown) composition* differing only in that one was ground in a three-roller mill with loose mill setting and the other with tight mill setting resulting in larger and smaller particle sizes, respectively. The panels which were measured were made by applying wet films of the paints 0.003 in. thick over white and black undercoats. The particle sizes were not known. Electron microscope photographs

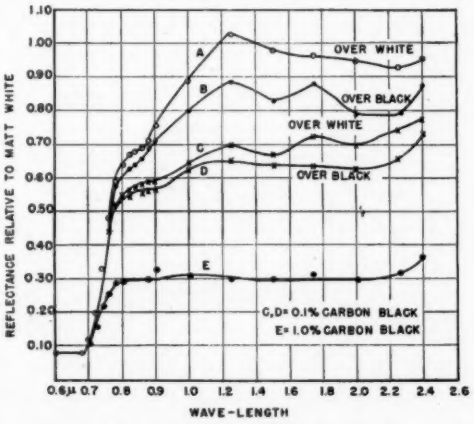

FIG. 6. Spectral reflectance of a stibnite pigment containing carbon black.

* These paints were made by the Paint Laboratory of the Philadelphia Naval Shipyard.

TABLE I. Contrast ratios $C=(R_w-R_B)/R_w$ derived from Fig. 5.

Wave-length	Fine particles	Coarse particles
0.6–0.760	0	0
0.78	.06	0
0.8	.12	.06
0.9	.16	.09
1.0	.22	.13
1.5	.25	.19
2.0	.23	.15
2.5	.21	.15

made by Mr. L. S. Birks showed that each pigment contained both large and small elongated particles with approximately rectangular surfaces, the proportion of small pigment crystals being much larger in the finely ground pigment.

The curves of Fig. 5 illustrate clearly that the larger particles are more effective scatterers of long wave-length light than smaller particles. Contract ratios $C=(R_w-R_B)/R_w$, where R_w is the reflectance over white and R_B, over black, are given in Table I to illustrate that above 0.76μ the coarser particles possessed higher hiding power than the smaller particles.

Effect of Absorption Coefficient

The effect of increasing the absorption coefficient of the layer of particles is illustrated by Fig. 6 which shows the reflectance of a stibnite paint over white and black surfaces (curves *A* and *B*), the reflectance of the same pigment containing 0.1 percent carbon black (curves *C* and *D*), and 1.0 percent carbon black (curve *E*). The reflectance of this aggregate of particles was the same over white and black undercoats and was constant at about 0.30 above 0.8μ. The addition of the carbon black did not appreciably affect the low (0.08) visible reflectance of the material since the absorption coefficient of the antimony sulfide already was high in that region.

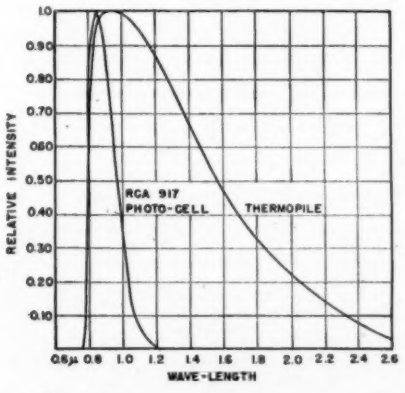

FIG. 7. Spectral-intensity distribution curves for photoelectric cell and thermocouple receivers with tungsten source and Wratten 87 filter.

INTEGRATED REFLECTANCE MEASUREMENTS

The foregoing sections have illustrated the type of information to be derived from spectral-reflectance measurements of pigments in the near infra-red. For many purposes, spectral measurements are not required and a single simple measurement in a known band of the spectrum will suffice. The instrument was adapted to such uses by employing a tungsten lamp of known spectral intensity distribution together with a chosen filter in place of a spectrometer. Either a thermocouple or a photoelectric cell was employed, depending upon the long wave-length limit desired, and in this method, rather insensitive instruments were required. Figure 7 shows the relative spectral distributions of intensity in such a system, employing a 100-watt Mazda projection lamp and a Wratten 87 filter, for both caesium oxide photoelectric cell and thermopile. The ordinates of the curve represent the product: $J_\lambda T_\lambda T_\lambda' S_\lambda$, where

J_λ is the spectral intensity of the tungsten lamp taken from the General Electric *Abstract Bulletin* (Nela Park, 1941).
T_λ is the spectral transmission of the Wratten filter.
T_λ' is the spectral transmission of the lens *L* of Fig. 1.
S_λ is the relative spectral sensitivity of the radiation detector.

It is evident by inspection of these curves that the use of a thermopile (or a lead sulphide photoconducting cell with peak sensitivity at 2μ, long wave-length threshold at 3μ) in place of the caesium oxide photo-cell extends the useful spectral range of the instrument well into the region in which the effects illustrated by Figs. 3 to 6 are pronounced. Indeed, when such an instrument is used for inspection and control during the formulation of materials of any required over-all optical characteristic in the near infra-red it has been found advantageous to employ a filter of even longer short wave-length limit as, say, Corning 2440 glass together with

the thermocouple in order to confine the effective spectral band of energy to wave-lengths above about 0.9μ. The effect of small variations in pigment characteristics which would escape observation with the usual photoelectric reflectometer may be pronounced when longer wave-lengths of light are used in making the measurement.

The use of such an integrating reflectometer in obtaining useful infra-red reflection factors is illustrated in Table II which shows the reflection factors obtained with the reflectometer for the paints described in detail in Figs. 5 and 6. The filter was Wratten 87. Measurements were made both with a caesium oxide photo-cell and a thermocouple. There is also included in Table II a column of reflection factors of the same paints for that portion of sunlight between 0.76 and 2.4μ. These factors were obtained by multiplying the reflection curves of Figs. 5 and 6 by the spectral intensity of sunlight received at the earth's surface through air mass 1, and taking the ratio of the integrated area under this curve to the area under the solar intensity curve between the same wave-length limits. The integrated reflection factors for sunlight are approximately equal to those measured with the tungsten source and thermocouple. Furthermore, the reflection factors obtained for the longer wave-length band of energy represent more truly the observed spectral reflectance factors about 1μ as may be seen by comparing Table II with Figs. 5 and 6.

TABLE II. Diffuse reflectance factors for the spectral bands 0.76–1.2μ (caesium oxide cell), 0.76–2.6μ (thermocouple) and integrated for sunlight 0.76–2.4μ.

Material	Photo-cell reflectance		Thermocouple reflectance		Integrated for sunlight	
	Over white	Over black	Over white	Over black	Over white	Over black
Coarse particles, Fig. 5	0.71	0.65	0.92	0.75	0.89	0.74
Fine particles, Fig. 5	0.73	0.62	0.90	0.66	0.86	0.64
Stibnite paint, Fig. 6	0.70	0.67	0.92	0.76	0.89	0.78
With 0.1 percent carbon black, Fig. 6	0.58	0.56	0.67	0.63	0.66	0.62
With 1.0 percent carbon black, Fig. 6	0.285	0.28	0.295	0.285	0.31	0.31

Infra-Red Image Tube Reflectometer

Frequently, measurements of reflectance with a caesium oxide photo-cell are useful, as in problems concerned with the photographic appearance of surfaces. Approximate measurements and rapid inspection of materials have often been made in this laboratory by viewing the surface with an infra-red image tube (RCA 1P25) which converts the infra-red image of the object viewed into a visible image. In such measurements a set of graded neutral gray cards was moved along next to the surface under examination until a photometric match in the visual images was obtained. The long wave-length threshold of this caesium oxide photoelectric surface is at 1.2μ. The short wave-length threshold was controlled by filters.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

The Limiting Colors Due to Ideal Absorption and Transmission Bands

I. H. GODLOVE
General Aniline and Film Corporation, Central Research Laboratory, Easton, Pennsylvania
(Received June 27, 1947)

Synthetic-organic chemists have developed to a high state the art of prediction of changes in absorption spectra with changes in molecular architecture. But they are not able to visualize the colors resulting from specific changes. The colorimetrist is frequently asked what spectra constitute ideal goals at which to shoot in a particular field of colors. Previously (see reference 7) the author had calculated the specifications of the colors caused by a V-shaped absorption band moving through the visible spectrum. In order to obtain a more detailed picture of the effects of changes in absorption spectra on the resulting colors, particularly in limiting cases, 100 selected ordinate calculations have now been made of the colors produced by idealized (rectangular) bands, of various widths, when moved across the spectrum.

The points representing the ideal colors in the 3-dimensional color solid lie on the same surface as do points found by MacAdam, but the new ones are of *much more direct interest to chemists*, for they relate to the two main and direct variables of the spectra, which are the chemists' current tools in precise studies of "color" and chemical structure.

The spectra are of two types with ideal cut-off (ordinates all zero or unity), for bands of various widths and wave-length centers. Munsell re-notations of all the colors have also been computed, and they have been compared in saturation with the excellent TCCA silk colors.

As previously pointed out, no true greens can be made with a single absorption band; and no magentas with a single transmission band. Saturation of dyes increases with width of absorption or narrowness of transmission bands; but a "law of diminishing returns" operates, there being a point beyond which it is unprofitable to go. A 5-mμ shift in the yellow has more effect, e.g., on the color than a 50-mμ shift in the blue-green. In two spectral-transmission regions, it makes little difference whether very wide or very narrow bands are used; in others the reverse is true. Saturation maxima were found at 3 absorption wave-length positions, and for transmission at 400–450 mμ. Many previous correlations of spectra and "color" are invalidated because of the "crowding" found of some hues into narrow wave-length regions, and great spread in others. In practice, yellows approach close to to the ideal maximum possible saturation, greens are very far away indeed; the spectral conditions for better greens are outlined.

INTRODUCTION

SYNTHETIC-ORGANIC chemists have developed to a high degree the art of prediction of changes in color and absorption spectra, which they think of as one and the same, with particular changes in chemical molecular structure.

But these predictions are likely to be in error in several respects, and fail to go far enough in specific ways to be described in this paper. They may proceed in a purely empirical manner to correlate color changes in dyed skeins with changes in chemical structure. This method is direct, simple, and readily understood; but it has disadvantages. It is difficult in this way to record the color changes in simple terms capable of precise graphical or numerical analysis; further, it is not easily possible to relate the color and structure changes to theoretical physical principles.

As a second method of procedure, chemists have with increasing frequency resorted to correlations of chemical structures with absorption spectra, which are physical properties of the dyes, usually thinking of the spectra as practically equivalent to or as determining the "color." This method of work has great advantages in the direction of graphical and numerical analysis and permits the application of a number of principles of modern physical theory, such as applications of electronic and resonance theories and theories of light absorption and transformation. But even when this method is used, although the proximate interest is absorption spectra, the ultimate interest is color perception; that is, how the color looks. It is notorious that no very simple relations exist between color perception and absorption spectra, hence no simple correlations of color and chemical structure. It is in fact common practice, when proceeding by the second method, to stop short at spectra, not taking the final step of proceeding to color perception.

* Presented at the meeting of the Optical Society of America, New York, February 22, 1947.

The relation between the necessary steps and ways of proceeding are made clear in Fig. 1.** At the center of the triangle we have indicated a colorant (dyestuff or pigment); at the corners we indicate the corresponding chemistry (lower left), physics (lower right), and color perception (or psychology, at the top). By chemical reactions we arrive at the classical chemical structure; by physical measurements we obtain physical data such as absorption spectra; and by direct observation the color perception. The sides of the triangle indicate the general types of correlations between the chemistry, the physics, and the color perception. The left leg embraces the older chemical correlations (especially those involved in the Witt theory). This is the chemist's direct method.

But we can go indirectly from color to chemical structure, and *vice versa*, through spectra. At the bottom are the newer physico-chemical studies, interpretations of spectra in terms of the electronic and resonance theories, and quantum-mechanical calculations. But the chemist's ultimate goal is color through spectra. Unfortunately, this has not been sufficiently well recognized, and the chemist stops short at spectra. But the right leg has been developed to a great science by members of this Society, and can be used to complete the journey involved in going indirectly from chemical structures to color perception, through spectra.

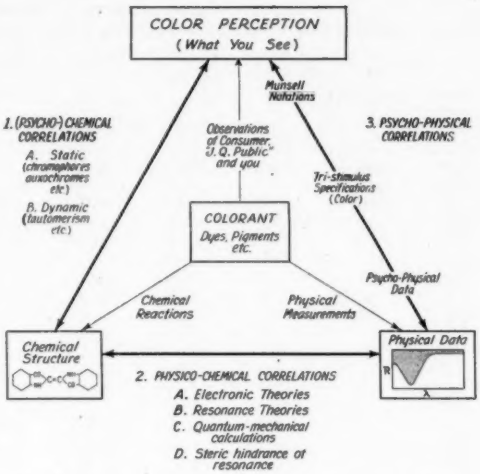

FIG. 1. Interrelations of color, colorant, and their chemistry and physics.

CALCULATION OF MAXIMALLY SATURATED COLORS

In recent years, chemists have been prone to an increasing degree to proceed through spectra, because this method permits the application of precision methods of measurement, hence greater utilization of numerical, mathematical, and graphical analysis.

A particularly useful idea in the theory of the relation of spectra, on the one hand to chemical structure, and on the other hand to color perception, is the *concept* of optimum colors of maximal saturation. Ostwald, Schrödinger, and recently MacAdam have made calculations of the colors of maximum possible saturation for a number of given hues and lightnesses and have developed the theory of the shapes of the absorption spectra necessary to produce such "best" colors. The points whose specifications MacAdam has calculated[1] lie on the outer bounding surface of the surface-color solid. The computations of this paper give the specifications of a great many points on the same surface, using the Ostwald-Schrödinger-[2,3]MacAdam theorems. But the new points are ones of *much more direct interest to chemists*, for they relate to the two main and direct variables of the absorption spectra, which are their current tools in precise studies of "color" and constitution. The present computations enable one to predict at once the colors resulting from absorption or transmission bands of given widths and given wave-length positions.

A frequent question asked of the colorimetrist by the chemist is of the nature: "What is my ideal goal in the absorption curves of synthesized dyes or pigments to achieve maximal saturation?" Related questions are: "In what direction do I need to shift the absorption to get a redder or a greener color, or a brighter one?" Or, how much broader or narrower should the absorption

** The original idea of the triangle of Fig. 1 was due to Dr. Warren F. Busse. The exact form and details were elaborated by the writer.

[1] MacAdam, J. Opt. Soc. Am. **25**, 249, 361 (1935).
[2] Ostwald, Kön. Sächs, Ges. d. Wiss. Abh. d. Math.-Phys. **34**, 471 (1917); Physik. Zeits. **17**, 328 (1916).
[3] Schrödinger, "Theorie der Pigmente von grösster Leuchkraft," Ann. d. Physik **62**, 603 (1920).

TABLE I. Colors produced by absorption bands of various positions and widths.

Center mμ	Reflectance Y	Trilinear (I.C.I.) coordinates x	y	Munsell notations V	H	C	Nearest TCCA color (but of lower chroma) name	C
1	2	3	4	5	6	7	8	9
Section 1. Transmission band of 30-mμ width								
390	.00	.0000	.0000	0.00	?	?		
410	.00	.1917	.0000	0.00	10−*PB*(?)	48 (?)		
430	.00	.1802	.0000	0.00	8 *PB*(?)	48 (?)		
450	.02	.1477	.0274	1.49	7−*PB*	33.0	U.S. Army ultramarine blue	11.8
460	.03	.1460	.0447	1.95	6.5 *PB*	26.6	U.S. Army cobalt blue	5.2
470	.03	.1281	.0560	1.95	6−*PB*	21.0	Navy 2	3.2
480	.05	.1020	.1300	2.61	9+*B*	12.8	Electric	4.6
490	.07	.0381	.2724	3.10	6 *BG*	20.4		
500	.11	.0000	.5085	3.84	8.5 *G*	30.0(?)		
510	.14	.0497	.7105	4.29	3+*G*	27.2	Primitive green	8.9
520	.19	.0800	.7754	4.91	2 *G*	28.3	Emerald	8.2
530	.23	.1686	.7908	5.34	1−*G*	24.5	Tarragon	3.4
540	.27	.2385	.7296	5.73	0.5 *G*	21.3		
550	.29	.3213	.6787	5.91	8.5 *GY*	18.9	Mintleaf	7.6
560	.29	.3783	.6217	5.91	6.5 *GY*	16.2	Spring green	7.2
570	.27	.4441	.5559	5.73	2 *GY*	13.7	Limepeel	6.0
580	.24	.5053	.4947	5.44	3.5 *Y*	13.5	Nugget gold	9.3
590	.20	.5696	.4304	5.03	5.5 *YR*	15.1	Burnished straw	5.5
600	.16	.6144	.3856	4.55	1.5 *YR*	15.8	Burnt orange	9.2
610	.12	.6713	.3287	4.00	9−*R*	18.3	Terra cotta	7.0
630	.07	.7041	.2957	3.10	8.5 *R*	17.0	U.S. Army brick red (?)	7.5
650	.03	.7235	.2768	1.95	10 *R*(?)	13.8		
670	.01	.7463	.2537	0.84	2.5 *YR*(?)	10.8		
690	.00	1.0000	.0000	0.00	?	?		
710	.00	—	—	0.00	?	?		
Section 2. Transmission band of 70-mμ width								
390	.00	.1917	.0000	0.00	10 *PB*(?)	48 (?)		
410	.00	.1718	.0000	0.00	7.5 *PB*(?)	48 (?)		
430	.02	.1563	.0199	1.49	7 *PB*(?)	37 (?)	U.S. Army cobalt blue (?)	5.2
450	.04	.1537	.0314	2.31	7 *PB*	38	U.S. Army ultramarine blue	11.8
460	.07	.1372	.0544	3.10	6.5 *PB*	29.8	Yale blue	8.8
470	.10	.1259	.0856	3.68	5.5 *PB*	21.6	Oriental blue (?)	11.1
480	.15	.1076	.1497	4.43	8 *B*	17.3	Electric	4.6
490	.20	.0954	.2434	5.03	10−*BG*	17.7	Parrot blue	6.5
500	.27	.0861	.3952	5.73	2+*BG*	24.4		
510	.35	.0947	.5632	6.41	5−*G*	29.4	Scarab green	9.3
520	.44	.1494	.6706	7.06	2.5+*G*	28.0		
530	.51	.2054	.7122	7.52	1+*G*	27.9		
540	.57	.2714	.6860	7.88	0+*G*	24.3	Crayon green	4.2
550	.60	.3390	.6483	8.05	8.5−*GY*	21.0		
560	.61	.3868	.6016	8.11	6.5−*GY*	19.2		
570	.58	.4479	.5521	7.94	1.5 *GY*	18.8		
580	.53	.4950	.5049	7.64	4.5 *Y*	19.0	Lemon yellow	8.7
590	.46	.5441	.4559	7.20	7.5+*YR*	20.2	U.S. Army yellow	13.2
600	.39	.5803	.4197	6.71	4 *YR*	20.4	Princeton orange	12.5
610	.31	.6219	.3781	6.08	0+*YR*	20.1	U.S. Army orange	13.2
630	.18	.6742	.3257	4.80	8.5 *R*	21.5	Paprika	14.3
650	.09	.7056	.2944	3.50	8 *R*	19.5	(Official U.S. flag red)?	9.6
670	.03	.7465	.2538	1.95	0+*YR*	15.6		
690	.01	.7464	.2538	0.84	2+*YR*	10.8		
710	.00	1.0000	.0000	0.00	?	?		
Section 3. Transmission band of 100-mμ width								
390	.00	.1771	.0000	0.00	7.5 *PB*(?)	48 (?)		
410	.01	.1653	.0112	0.84	7.5−*PB*(?)	39 (?)		
430	.03	.1498	.0241	1.95	7 *PB*	39.5	Royal blue	8.4
450	.08	.1391	.0567	3.31	6.5 *PB*	30.0	Lustre blue	7.9
460	.12	.1341	.0821	4.00	6 *PB*	24.0	Hydrangea blue	5.1
470	.18	.1252	.1209	4.80	3.5−*PB*	19.6	Bluebird(?)	6.4
480	.25	.1204	.1705	5.54	7.5 *B*	17.2	U.S. Army blue (?)	3.4
490	.34	.1217	.2484	6.33	1+*B*	17.9	Blue turquoise (?)	5.6
500	.42	.1279	.3424	6.93	5 *BG*	21.3		
510	.52	.1468	.4579	7.58	8 *G*	23.4		
520	.61	.1893	.5609	8.11	3.5 *G*	24.0		
530	.68	.2477	.6137	8.48	1+*G*	22.4		

TABLE I.—*Continued.*

Center mμ	Reflectance Y	Trilinear (I.C.I.) coordinates x	y	Munsell notations V	H	C	Nearest TCCA color (but of lower chroma) name	C
1	2	3	4	5	6	7	8	9
540	.74	.3002	.6295	8.78	9.5 *GY*	21.6	Crayon green (?)	4.2
550	.77	.3506	.6119	8.92	8 *GY*	19.1		
560	.77	.4002	.5820	8.92	5.5+*GY*	18.3		
570	.75	.4438	.5476	8.83	1.5+*GY*	18.8		
580	.71	.4843	.5158	8.63	5.5 *Y*	20.4	Lemon yellow (?)	8.7
590	.63	.5178	.4822	8.22	1.5− *Y*	20.6	Spanish yellow (?)	12.2
600	.56	.5471	.4529	7.82	7.5 *YR*	21.5	U.S. Army yellow	13.2
610	.46	.5845	.4155	7.20	3.5− *YR*	22.1	Princeton organge	12.5
630	.29	.6418	.3582	5.91	9.5 *R*	21.8	Salmon	10.7
650	.15	.6959	.3042	4.43	8− *R*	22.7	Paprika	14.3
670	.07	.7159	.2840	3.10	8+*R* (?)	18.5	(Official U.S. flag red)?	9.6
690	.03	.6958	.3042	1.95	1 *YR* (?)	12.1		
710	.01	.6623	.3377	0.84	3.5 *YR*(?)	6.4		
				Section 4. Transmission band of 130-mμ width				
390	.01	.1669	.0131	0.84	7+*PB*	38 (?)		
410	.03	.1572	.0253	1.95	7.5− *PB*	38.5(?)	National flag blue	2.5
430	.07	.1412	.0504	3.10	6.5 *PB*	32.2	Yale blue	8.8
450	.15	.1294	.0990	4.43	5+*PB*	21.7	Hydrangea blue	5.1
460	.21	.1299	.1325	5.13	2.5+*PB*	19.4	Bluebird (?)	6.4
470	.29	.1279	.1719	5.91	8.5 *B*	17.2	U.S. Army light blue	3.4
480	.38	.1375	.2132	6.64	5*B*	16.4		
490	.48	.1481	.2590	7.33	2− *B*	16.1	} Turquoise (?)	3.7
500	.58	.1651	.3151	7.94	7.5 *BG*	16.7		
510	.67	.1895	.3808	8.43	2+*BG*	16.7		
520	.74	.2279	.4411	8.78	5+*G*	15.8		
530	.80	.2765	.5014	9.06	2− *G*	15.3		
540	.84	.3220	.5410	9.24	9.5 *GY*	15.4		
550	.87	.3653	.5588	9.37	7+*GY*	16.1		
560	.87	.4105	.5520	9.37	4.5 *GY*	16.4		
570	.87	.4346	.5432	9.37	2+*GY*	17.4		
580	.83	.4662	.5263	9.20	8+*Y*	19.8		
590	.78	.4848	.5075	8.97	5− *Y*	20.3	Lemon yellow	8.7
600	.70	.5192	.4808	8.58	1+*Y*	21.2	Spanish yellow (?)	12.2
610	.62	.5392	.4608	8.16	8+*YR*	21.6	U.S. Army yellow	13.2
630	.42	.5990	.4010	6.93	2 *YR*	21.8		
650	.25	.6531	.3469	5.54	9 *R*	21.8	U.S. Army orange	13.2
670	.13	.6935	.3065	4.15	8 *R*	21.1		
690	.06	.6958	.3042	2.87	9+*R*(?)	15.4	U.S. Army brick red	7.5
710	.02	.7102	.2898	1.49	1.5*R*(?)	11.1		
				Section 5. Transmission band of 170-mμ width				
390	.03	.1572	.0253	1.95	7.5− *PB*	38.5(?)	Sapphire blue	10.0
410	.07	.1412	.0504	3.10	6.5+*PB*	32.2	U.S. Army ultramarine blue	11.8
430	.15	.1294	.0990	4.43	5+*PB*	21.7	Hydrangea blue	5.1
450	.29	.1279	.1719	5.91	8.5 *B*	17.2	U.S. Army light blue	3.4
460	.38	.1357	.2104	6.64	5 *B*	16.6		
470	.48	.1504	.2455	7.33	3+*B*	15.8		
480	.58	.1669	.2743	7.94	1.5 *B*	14.7		
490	.67	.1855	.2948	8.43	0+*B*	13.3		
500	.75	.2109	.3103	8.83	9.5− *BG*	10.8		
510	.82	.2317	.3284	9.15	7− *BG*	9.1		
520	.87	.2586	.3531	9.37	1 *BG*	7.3		
530	.91	.2888	.3885	9.54	3+*G*	6.9		
540	.93	.3252	.4226	9.62	9 *GY*	7.5		
550	.95	.3605	.4658	9.70	6.5 *GY*	9.3		
560	.95	.3906	.4980	9.70	5 *GY*	11.8		
570	.95	.4141	.5211	9.70	3 *GY*	14.3		
580	.93	.4392	.5273	9.62	1 *GY*	16.6		
590	.90	.4529	.5263	9.50	10− *Y*	17.8		
600	.85	.4765	.5164	9.29	6 *Y*	20.1		
610	.79	.4913	.5012	9.02	4 *Y*	20.1	Jasmine	7.7
630	.62	.5425	.4575	8.16	8− *YR*	21.7	U.S. Army yellow	13.2
650	.42	.5990	.4010	6.93	2 *YR*	21.8	U.S. Army golden orange	11.4
670	.25	.6531	.3469	5.54	9 *R*	21.8	U.S. Army orange	13.2
690	.13	.6935	.3065	4.15	8 *R*	21.1		
710	.06	.6958	.3042	2.87	9+*R*(?)	15.4	U.S. Army brick red	7.5

TABLE I.—*Continued.*

Center mμ	Reflectance Y	Trilinear (I.C.I.) coordinates x	y	Munsell notations V	H	C	Nearest TCCA color (but of lower chroma). name	C
1	2	3	4	5	6	7	8	9
				Section 6. Transmission band of 200-mμ width				
390	.05	.1458	.0372	2.61	7−*PB*	35.9	U.S. Army ultramarine blue	11.8
410	.12	.1331	.0814	4.00	6 *PB*	24.0	Hydrangea blue	5.1
430	.25	.1258	.1527	5.54	0.5−*PB*	18.2	U.S. Army light blue	3.4
450	.43	.1411	.2292	7.00	4 *B*	16.2		
460	.53	.1549	.2617	7.64	2 *B*	16.1		
470	.63	.1743	.2872	8.22	0.5 *B*	14.7		
480	.71	.1959	.3019	8.63	10−*BG*	12.3		
490	.79	.2178	.3135	9.02	9 *BG*	9.9		
500	.85	.2388	.3185	9.29	8 *BG*	7.6		
510	.89	.2605	.3196	9.46	7.5+*BG*	5.2		
520	.93	.2778	.3254	9.62	6 *BG*	3.8		
530	.96	.2966	.3377	9.74	7 *G*	2.7		
540	.97	.3151	.3626	9.78	10−*GY*	3.4		
550	.98	.3383	.3979	9.82	6.5 *GY*	5.1		
560	.98	.3629	.4370	9.82	5+*GY*	7.4		
570	.98	.3882	.4790	9.82	4+*GY*	10.4		
580	.97	.4141	.5058	9.78	2.5−*GY*	12.1		
590	.95	.4316	.5228	9.70	1.5 *GY*	15.5		
600	.92	.4478	.5252	9.58	0+*GY*	16.9		
610	.88	.4647	.5213	9.42	8−*Y*	18.5		
630	.75	.5080	.4920	8.83	2.5+*Y*	22.4	Maize (?)	6.8
650	.57	.5567	.4433	7.88	6 *YR*	21.9	U.S. Army yellow (?)	13.2
670	.37	.6178	.3822	6.56	0.5 *YR*	22.5	Dusty pink	3.9
690	.21	.6726	.3274	5.13	8.5 *R*	22.8	U.S. Army orange (?)	13.2
710	.11	.6985	.3015	3.84	8+*R*	20.2	(Official U.S. flag red)?	9.6
				Section 7. Absorption band of 30-mμ width				
390	1.00	.3101	.3163	9.90	*N*	0.0		
410	1.00	.3141	.3269	9.90	6.5+*GY*	0.6		
430	1.00	.3339	.3743	9.90	5.5+*GY*	3.4		
450	.98	.3589	.4031	9.82	2.5 *GY*	5.2		
460	.97	.3543	.3895	9.78	2−*GY*	4.3		
470	.97	.3472	.3694	9.78	8.5+*Y*	3.3	Jasmine (?)	7.7
480	.95	.3389	.3421	9.70	9+*YR*	2.1		
490	.93	.3342	.3202	9.62	5*R*	2.9		
500	.89	.3329	.3022	9.46	4.5 *RP*	4.8		
510	.86	.3274	.2901	9.33	1.5 *RP*	5.8		
520	.81	.3294	.2777	9.11	1 *RP*	7.4		
530	.77	.3244	.2682	8.92	9.5+*P*	8.4		
540	.73	.3196	.2615	8.73	9 *P*	9.0		
550	.71	.3084	.2597	8.63	7 *P*	8.5		
560	.71	.2983	.2634	8.63	6−*P*	7.4		
570	.73	.2858	.2728	8.73	1 *P*	5.1		
580	.76	.2747	.2840	8.87	4.5 *PB*	3.8		
590	.80	.2654	.2966	9.06	6+*B*	4.0		
600	.84	.2641	.3058	9.24	2.5+*B*	4.3		
610	.88	.2629	.3147	9.42	9.5+*BG*	4.6		
630	.93	.2782	.3180	9.62	8.5+*BG*	3.3		
650	.97	.2954	.3177	9.78	8.5 *BG*	1.6		
670	.99	.3046	.3171	9.86	7.5 *BG*	0.6		
690	1.00	.3057	.3183	9.90	5 *BG*	0.5		
710	1.00	.3101	.3163	9.90	*N*	0.0		
				Section 8. Absorption band of 50-mμ width				
390	1.00	.3124	.3187	9.90	8 *YR*(?)	0.1		
410	1.00	.3225	.3462	9.90	6+*GY*	1.8		
430	.99	.3557	.4128	9.86	5+*GY*	6.0		
450	.97	.3914	.4666	9.78	3−*GY*	9.7		
460	.96	.3902	.4495	9.74	1.5 *GY*	8.3		
470	.93	.3815	.4067	9.62	6−*Y*	5.8	Jasmine (?)	7.7
480	.91	.3672	.3665	9.54	10 *YR*	4.1		
490	.87	.3552	.3283	9.37	7+*R*	4.0		
500	.82	.3480	.2970	9.15	5+*RP*	6.6		
510	.75	.3464	.2704	8.83	2+*RP*	9.7		
520	.69	.3395	.2515	8.53	0.5+*RP*	12.0		

TABLE I.—*Continued.*

Center mμ	Reflectance Y	Trilinear (I.C.I.) coordinates x	y	Munsell notations V	H	C	Nearest TCCA color (but of lower chroma) name	C
1	2	3	4	5	6	7	8	9
530	.62	.3332	.2341	8.16	10−*P*	14.3		
540	.57	.3213	.2224	7.88	7.5 *P*	14.1		
550	.54	.3091	.2183	7.70	7.5−*P*	14.7	Orchid (?)	5.8
560	.45	.2851	.2243	7.70	5 *P*	12.0		
570	.56	.2649	.2364	7.82	0+*P*	8.9		
580	.61	.2472	.2564	8.11	4+*PB*	7.0		
590	.67	.2350	.2769	8.43	7+*B*	7.2		
600	.73	.2262	.2955	8.73	2+*B*	8.4		
610	.78	.2337	.3048	8.97	1 *B*	7.5		
630	.88	.2577	.3169	9.42	9−*BG*	5.4		
650	.94	.2868	.3161	9.66	0+*B*	2.5		
670	.98	.3012	.3169	9.82	9.5 *BG*	0.9		
690	1.00	.3080	.3173	9.90	5 *BG*	0.5		
710	1.00	.3101	.3163	9.90	*N*	0.0		
				Section 9. Absorption band of 70-mμ width				
390	1.00	.3141	.3269	9.90	6.5+*GY*	0.6		
410	1.00	.3369	.3776	9.90	5.5+*GY*	3.7		
430	.98	.3816	.4541	9.82	3.5−*GY*	8.4		
450	.96	.4158	.5090	9.74	2.5−*GY*	13.4		
460	.93	.4286	.4958	9.62	10 *Y*	12.8		
470	.90	.4181	.4515	9.50	6.5−*Y*	9.8	Jasmine (?)	7.7
480	.85	.4040	.3936	9.29	10−*YR*	6.8		
490	.80	.3855	.3419	9.06	9.5 *R*	5.9		
500	.73	.3718	.2945	8.73	7−*RP*	9.1		
510	.65	.3628	.2559	8.32	3−*RP*	13.3		
520	.56	.3522	.2235	7.82	0.5+*RP*	16.9		
530	.49	.3408	.2004	7.39	10 *P*	19.9		
540	.43	.3239	.1845	7.00	8.5 *P*	21.3		
550	.40	.2981	.1789	6.78	7−*P*	20.0	Orchid	5.8
560	.39	.2739	.1816	6.71	5 *P*	17.6		
570	.42	.2415	.1990	6.93	10−*PB*	12.6		
580	.47	.2182	.2225	7.26	4−*PB*	10.3		
590	.54	.2004	.2509	7.70	7−*B*	10.2		
600	.61	.1976	.2733	8.11	3.5 *B*	11.0		
610	.69	.2010	.2946	8.53	1−*B*	11.7		
630	.82	.2330	.3143	9.15	8.5 *BG*	8.0		
650	.91	.2678	.3186	9.54	8−*BG*	4.3		
670	.97	.2931	.3187	9.78	7.5 *BG*	1.8		
690	.99	.3046	.3171	9.86	7.5 *BG*	0.6		
710	1.00	.3080	.3173	9.90	5 *BG*	0.5		
				Section 10. Absorption band of 100-mμ width				
390	1.00	.3284	.3605	9.90	6+*GY*	2.6		
410	.99	.3667	.4357	9.86	5−*GY*	7.5		
430	.97	.4141	.5059	9.78	2.5−*GY*	12.1		
450	.92	.4478	.5252	9.58	0+*GY*	16.9		
460	.88	.4614	.5177	9.42	8 *Y*	17.6		
470	.82	.4746	.4901	9.15	4+*Y*	16.0	Jasmine	7.7
480	.75	.4741	.4423	8.83	10 *YR*	12.8		
490	.66	.4539	.3682	8.38	2−*YR*	10.2	Salmon pink (??)	7.7
500	.58	.4256	.2997	7.94	10 *RP*	12.7	Vassar rose (??)	8.1
510	.48	.4016	.2369	7.33	4 *RP*	18.4		
520	.39	.3734	.1880	6.71	1.5+*RP*	23.9		
530	.32	.3437	.1558	6.16	10−*P*	27.5	Crocus	6.4
540	.26	.3159	.1309	5.64	8 *P*	29.1		
550	.23	.2833	.1208	5.34	6+*P*	28.4	Lavender	4.8
560	.23	.2453	.1251	5.34	3 *P*	24.7		
570	.25	.2079	.1395	5.54	9.5 *PB*	19.3	Periwinkle (?)	6.8
580	.29	.1803	.1616	5.91	5.5+*PB*	15.5		
590	.37	.1638	.1994	6.56	10−*B*	14.7	Sky blue	3.4
600	.44	.1579	.2286	7.06	5 *B*	14.4		
610	.54	.1622	.2628	7.70	2+*B*	15.0		
620	.63	.1780	.2859	8.22	1−*B*	13.9		
630	.71	.1959	.3019	8.63	10−*BG*	12.3		
650	.85	.2388	.3185	9.29	8 *BG*	7.6		
670	.93	.2758	.3190	9.62	8+*BG*	3.7		

TABLE I.—*Continued.*

Center mμ	Reflectance Y	Trilinear (I.C.I.) coordinates x	y	Munsell notations V	H	C	Nearest TCCA color (but of lower chroma) name	C
1	2	3	4	5	6	7	8	9
690	.97	.2977	.3167	9.78	9.5+*BG*	1.4		
710	.99	.3068	.3161	9.86	10+*BG*	0.4		
				Section 11. Absorption band of 130-mμ width				
390	.99	.3557	.4128	9.86	5+*GY*	6.0		
410	.97	.4017	.4907	9.78	2.5 *GY*	11.6		
430	.93	.4423	.5244	9.62	0.5 *GY*	16.4		
450	.85	.4765	.5164	9.29	6 *Y*	20.1		
460	.79	.4913	.5012	9.02	4 *Y*	20.1	Lemon yellow (?)	8.7
470	.71	.5185	.4815	8.63	1 *Y*	22.6	Maize	6.8
480	.62	.5332	.4496	8.16	7.5 *YR*	19.1	U.S. Army yellow (?)	13.2
490	.52	.5394	.3974	7.58	2.5 *YR*	15.6	Peach	5.7
500	.42	.5121	.3180	6.93	5−*R*	16.1	Melon pink	11.5
510	.33	.4614	.2353	6.25	6+*RP*	21.4	Ashes of rose	5.9
520	.26	.4030	.1752	5.64	2.5 *RP*	26.8	Fuchsia (?)	17.2
530	.20	.3443	.1277	5.03	10−*P*	30.2		
540	.16	.2986	.0994	4.55	7.5 *P*	32.1	Lilac	6.4
550	.13	.2566	.0810	4.15	5 *P*	33.0	Violet (??)	11.0
560	.13	.2102	.0820	4.15	1+*P*	30.1	Parma violet	10.5
570	.13	.1822	.0833	4.15	9 *PB*	27.4	Periwinkle (?)	8.7
580	.17	.1547	.1073	4.68	6.5−*PB*	21.3		
590	.22	.1448	.1354	5.24	4.5+*PB*	18.4	Copenhagen	4.8
600	.30	.1321	.1758	6.00	8+*B*	16.1	U.S. Army light blue (?)	3.4
610	.38	.1404	.2092	6.64	5 *B*	16.1		
630	.58	.1669	.2743	7.94	1.5 *B*	14.7		
650	.75	.2089	.3072	8.83	10−*BG*	10.8		
670	.87	.2507	.3178	9.37	7.5+*BG*	6.2		
690	.94	.2844	.3171	9.66	9−*BG*	2.6		
710	.98	.3012	.3169	9.82	9.5 *BG*	0.9		
				Section 12. Absorption band of 200-mμ width				
350	.99	.3473	.3941	9.86	5 *GY*	4.6		
390	.95	.4317	.5228	9.70	1.5 *GY*	14.4		
410	.88	.4647	.5214	9.42	8−*Y*	18.5		
430	.75	.5080	.4920	8.83	2.5+*Y*	22.4	Popcorn	7.2
450	.57	.5567	.4433	7.88	6 *YR*	21.9	Apricot (?)	6.1
460	.47	.5865	.4135	7.26	3 *YR*	22.2	Honeydew	8.3
470	.37	.6178	.3822	6.56	0.5 *YR*	22.5	Dusty pink (?)	3.9
480	.29	.6418	.3582	5.91	9.5 *R*	21.8	U.S. Army orange	13.2
490	.21	.6726	.3274	5.13	8.5 *R*	22.8	Flame red (?)	12.7
500	.15	.6959	.3042	4.43	8−*R*	22.7	Tangerine	11.9
510	.11	.6767	.2920	3.84	7−*R*	19.2	Pimento	14.1
520	.07	.6137	.2306	3.10	2−*R*	18.0	American beauty (?)	8.8
530	.04	.4303	.1254	2.31	4.5−*RP*	21.0	Plum (?)	4.9
540	.03	.2823	.0617	1.95	7 *P*	25.7	Dahlia purple (?)	10.2
550	.02	.2105	.0286	1.49	1.5 *P*	33.5	Purple	6.9
560	.02	.1814	.0218	1.49	8.5 *PB*	36.5	(Official U.S. flag blue)?	1.8
570	.02	.1669	.0179	1.49	7.5 *PB*	39.0	Sapphire blue (?)	10.0
580	.03	.1498	.0241	1.95	7 *PB*	39.5	Royal blue	8.4
590	.05	.1458	.0372	2.61	7−*PB*	35.9	U.S. Army ultramarine blue	11.8
600	.08	.1391	.0567	3.31	6.5 *PB*	30.0	Lustre blue	7.9
610	.12	.1331	.0814	4.00	6 *PB*	24.0	Hydrangea blue	5.1
630	.25	.1258	.1527	5.54	0.5−*PB*	18.2	Saxe blue	4.6
650	.43	.1411	.2292	7.00	4 *B*	16.2		
670	.63	.1743	.2872	8.22	0.5 *B*	14.7		
690	.79	.2178	.3135	9.02	9 *BG*	9.9		
710	.89	.2594	.3182	9.46	8+*BG*	5.6		
750	.99	.3024	.3181	9.86	7 *BG*	0.9		
				Section 13. Absorption band of 300-mμ width				
330	.97	.4141	.5059	9.78	2.5−*GY*	12.1		
350	.92	.4478	.5252	9.58	0+*GY*	16.9		
370	.82	.4853	.5074	9.15	5−*Y*	20.3	Lemon yellow (?)	8.7
390	.66	.5303	.4697	8.38	9.5+*YR*	21.4	Spanish yellow	12.2
410	.47	.5865	.4135	7.26	3 *YR*	22.2	Peach	5.7
430	.29	.6418	.3582	5.91	9.5 *R*	21.8	U.S. Army orange	13.2
450	.15	.6959	.3042	4.43	8−*R*	22.7	Lacquer	9.5

TABLE I.—*Continued.*

Center mμ	Reflectance Y	Trilinear (I.C.I.) coordinates x	y	Munsell notations V	H	C	Nearest TCCA color (but of lower chroma) name	C
1	2	3	4	5	6	7	8	9
470	.07	.7159	.2840	3.10	8+*R*	18.5	(Official U.S. flag red)?	9.6
490	.03	.6960	.3043	1.95	1 *YR*(?)	12.1(?)		
510	.01	.6625	.3378	0.84	3.5 *YR*(?)	6.4(?)		
530	.00	—	—	0.00				
550	.00	—	—	0.00				
570	.00	.1719	.0000	0.00	7 *PB*(?)	44 (?)		
590	.00	.1771	.0000	0.00	7.5 *PB*(?)	48 (?)		
610	.01	.1653	.0112	0.84	7.5−*PB*(?)	39 (?)		
630	.03	.1498	.0241	1.95	7 *PB*	39.5	Royal blue	8.4
650	.08	.1391	.0567	3.31	6.5 *PB*	30.0	Lustre blue	7.9
670	.18	.1269	.1165	4.80	4 *PB*	19.7	Copenhagen	4.8
690	.34	.1340	.1936	6.33	7−*B*	17.0		
710	.53	.1549	.2617	7.64	2 *B*	16.1		
730	.71	.1959	.3019	8.63	10−*BG*	12.3		
750	.85	.2388	.3185	9.29	8 *BG*	7.6		

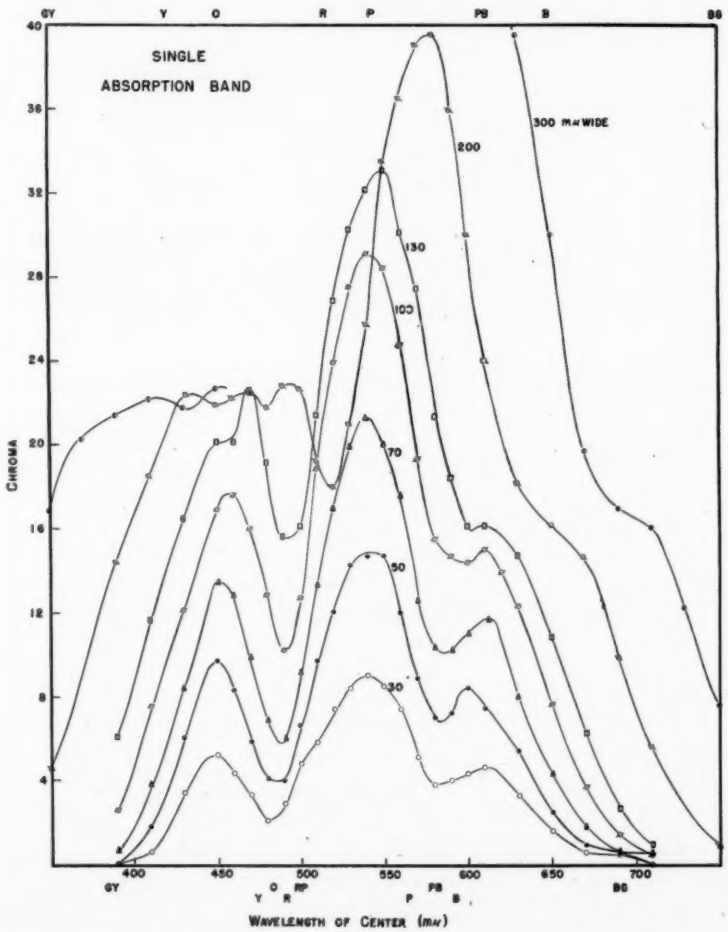

FIG. 2. Wave-length center of band *vs.* chroma for various band widths; single absorption bands.

bands be; or, will one band in the visible spectrum suffice, or are two needed.

Moreover, the chemist wishes to know, when he has prepared for a particular purpose a dye of some measured spectral-absorption characteristics, how this dye compares in saturation with the "best" or "brightest" color obtainable with the given hue and concentration.

THE OSTWALD-SCHRÖDINGER-MACADAM THEOREMS

According to the theories of these workers, the limiting or "best" colors, those of maximum possible saturations in the given hue (and concentration) range are those obtained by means of idealized absorption or transmission bands. these ideal bands have all absorption values, at every wave-length, equal either to 100 percent or to zero. That is, they have perfectly sharp transitions between the ranges of absorption and transmission (or reflection). They jump at once from zero to unit transmission at the transition wave-length.

One transmission band near the middle of the visible spectrum is the equivalent of two absorption bands, and *vice versa;* and MacAdam[1] and the writer have both considered these two cases, one and two absorption bands. Both the MacAdam colors and the new ones of this paper, which all lie on the outer bounding surface of

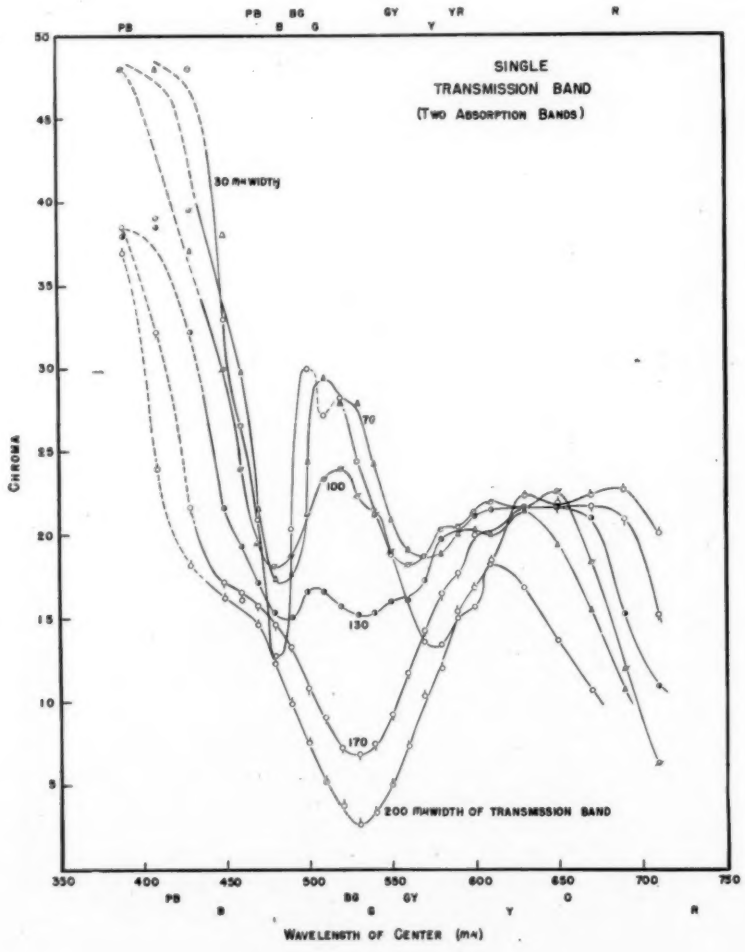

FIG. 3. Wave-length center of band *vs.* chroma for various band widths; single transmission bands (two absorption bands).

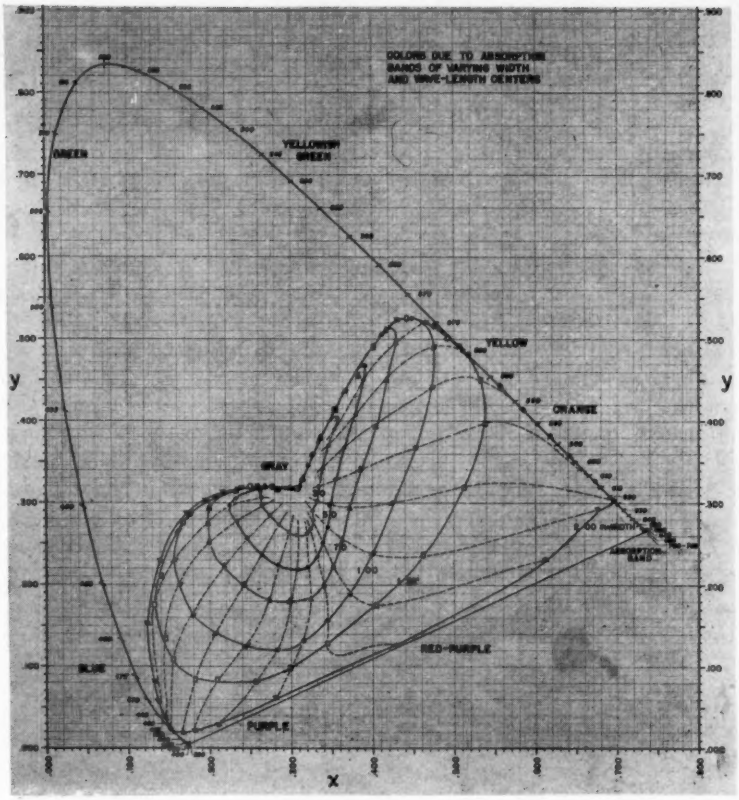

FIG. 4.

the surface-color solid (*but concentrated at different places* for the two sets) can be represented by projection on the I.C.I. chromaticity diagram. But on this projection the one-transmission-band colors all lie above, and the one-absorption-band colors all lie below a certain "alar" (wing-shaped) curve which passes through the daylight point at the center of the diagram.

MacAdam[1] stopped at an I.C.I. specification and did not convert these to color-perception terms. Nickerson and Newhall, however, converted the MacAdam specifications to specifications in terms of the Munsell color system. The colors of the present paper were also converted to Munsell specifications. In addition, for the new set of colors the TCCA color nearest in hue and lightness but of lower saturation (Munsell chroma) has been found and tabulated. This serves to bring out how far chemists have been able to go with present-day dyes in the approach to the maximal saturation possible with "ideal" absorption bands.

PROCEDURE USED IN MAKING CALCULATIONS

The methods used were extremely simple and rapid. To obtain the "1931 I.C.I."[4] or tristimulus specification, use was made of the selected-ordinate method of computation,[5] using 100 ordinates. The results which are found and tabulated are the coordinates x and y and the lightness Y.

To facilitate the computation a simple mode of operation was set up. Since all the ordinates for the limiting ("best") colors are either unity or zero, it was *only necessary to count the number or ordinates* having unit (not zero) value. To

[4] Judd, J. Opt. Soc. Am. **23**, 359 (1933).
[5] Hardy, *Handbook of Colorimetry* (Technology Press (of M.I.T.), Cambridge, Massachusetts, 1936); O.S.A. Colorimetry Committee, J. Opt. Soc. Am. **34**, 648 (1944)

make this process very rapid and practically "fool-proof," this was done in the following way.

The lists of selected ordinates, taken separately for X, Y, and Z from reference 5, were evenly spaced on long thin strips of cross-sectional paper. On a similar strip of cross-sectional paper were put the numbers 1 to 100. The number strip was merely brought along side of the wave-length strip for X, with the zero index at the short-wave limit of the absorption (or transmission) band under consideration. The number of wave-lengths or unit ordinates included between this zero position and the long-wave limit of the band was then immediately read off. The same operation was repeated for Y and for Z. The whole operation required only a few seconds.

As an example (the results being given in the first section of Table I) for a band of 30-mμ width centering at 600 mμ, the short- and long-wave limits are 585 mμ and 615 mμ. The selected ordinates for X between these limits are successively: 585.9, 587.0, 588.1, 589.2, 590.3, 591.4, 592.5, 593.6, 594.6, 595.7, 596.8, 597.9, 599.0, 600.1, 601.2, 602.3, 603.4, 604.5, 605.6, 606.8, 607.9, 609.1, 610.2, 611.4, 612.6, and 613.9 mμ. It is necessary only to note that there are 26 ordinates within the interval; and this is done mechanically in the way already indicated. In a similar way it is found that there are 16 ordinates selected for Y in the same interval, and none for Z. The sum of $26+16+0$ is equal to 42. 26 divided by 42 equals 0.6144, which is recorded in column 3 under "x"; and $16/42 = 0.3856$, which is "y" (column 4).

The Munsell specification is given in columns 5, 6, and 7, where the headings are "V" for "value" (lightness), "H" for "hue," and "C" for "chroma" (saturation). The letters for the hue notations are the initial letters of hue names. These specifications were read off di-

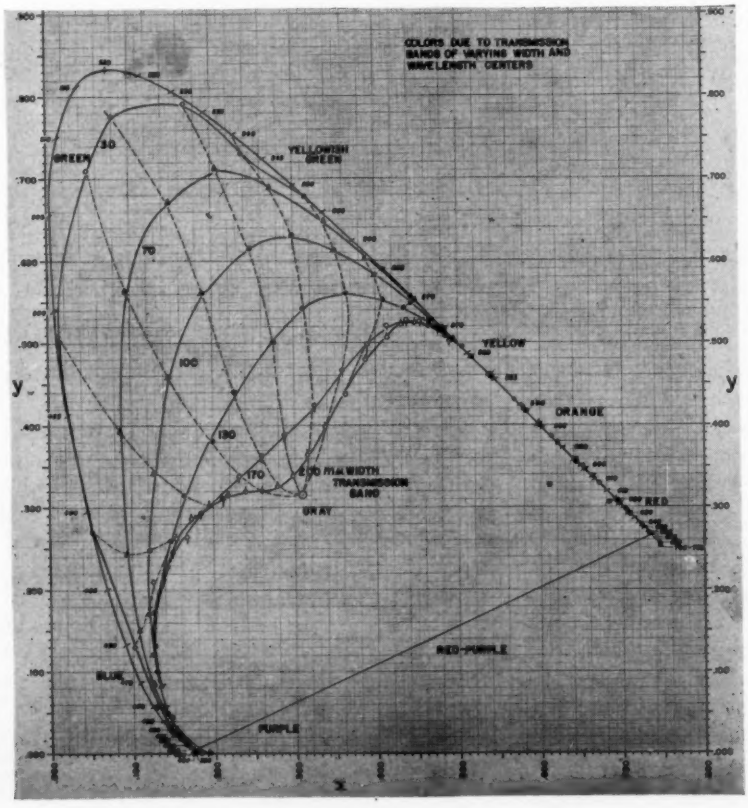

FIG. 5.

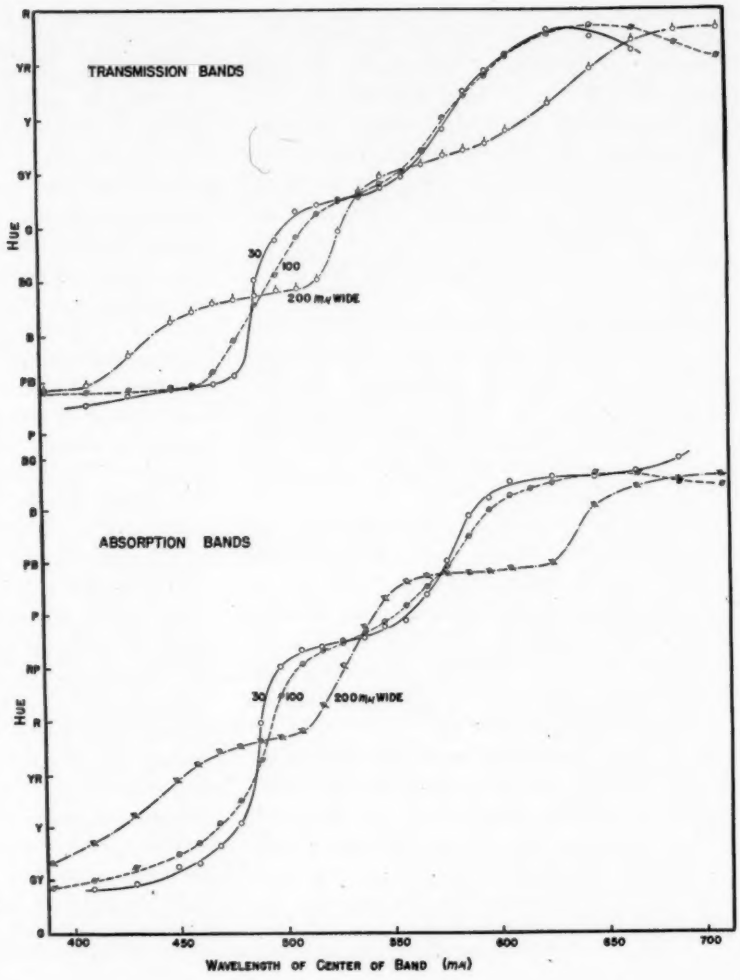

FIG. 6. Hue *vs.* wave-length centers of bands for three widths of transmission and absorption bands.

rectly from very large-scale reproductions of graphs given in the Newhall, Nickerson, and Judd reference.[6] These are the so-called "Munsell renotations."

In column 8 are written the names of the nearest colors, in the standards of the Textile Color Card Association of the U. S., corresponding to the calculated colors on the same horizontal line. The 1931 I.C.I. specifications of these were determined at the National Bureau of Standards. The hues and Munsell "values" of these colors are reasonably close to the calculated maximal ones, but the chromas (saturations) are the maximum obtainable in dyed or permanent pigment surfaces today. The chromas of these colors are given in column 9.

[6] Newhall, Nickerson, and Judd, J. Opt. Soc. Am. **33**, 385 (1943).

RESULTS

The results of this investigation may be conveniently summarized as follows:

As shown in Fig. 2 for *absorption* bands of 30–130-mμ width, there are peaks of saturation (chroma) at about 450, 540, and 610 mμ, corresponding to yellow, purple, and blue hues. The highest peaks are in the purples. For *transmission* bands of 30–200-mμ width (Fig. 3), saturations

are high for centers from 400–450 mμ, low for centers from 650 mμ to 700 mμ, and exhibit the following peculiarities in the region between these terminal ranges: Near 480 mμ (greenish blue) and 620 mμ (yellow), it makes little difference in the saturation whether a very narrow or a very wide band is used; from 500 to 550 mμ, however, there is an enormous variation of saturation with band width (3 chroma steps for 200-mμ width and 30 chroma steps for 30-mμ width).

Saturations are, of course, greatest with wide absorption bands and narrow transmission bands; but a "law of diminishing returns" appears to operate, for there is little advantage in widening the absorption bands beyond 200 mμ or narrowing the transmission to less than 30 mμ. (See Fig. 1.)

It is seen in Fig. 4 that on the color "mixture diagram according to the 1931 I.C.I. standard observer and coordinate system," all of the *single absorption band colors* fall on or below a certain "alar" (wing-shaped) path through the central gray point. All of the single-transmission-band (two-absorption-band) colors fall above the alar path (Fig. 5).

As a *single absorption* band (of any width from 10–300 mμ) is passed from the ultraviolet through the visible spectrum to the infra-red, the hues pass from green-yellow through yellow, red, purple, and blue to blue-green, *but true greens can never be obtained* in this way. See the bottom of Figs. 2 and 3.

With a *single transmission* band the hues pass from purple-blue through blue, green, and yellow to red. No purples or magentas are obtained.

It may be noted in Fig. 6 that the rate of change of hue with wave-length center is extremely irregular. For example, for a single narrow absorption band, the hues yellow, orange, red, and magenta are crowded together in a 30-mμ range of band centers from 470–500 mμ, while the green-yellows and the blue-greens are each spread out over a 100-mμ range. This fact makes of doubtful value many previously published correlations of "color" and chemical structure based on shift of absorption bands. For a 5-mμ change in the orange range may be more significant visually than a 50-mμ change in the blue-green, for example. (These facts are shown for 3 widths of both transmission and absorption bands in Fig. 6, where hue is plotted against wave-length center.)

As a single absorption or transmission band moves across the visible spectrum, the resultant colors give the symmetrical curves shown in Fig. 7 for the relation of lightness (Munsell "value") and wave-length center. For the absorption bands, the colors pass from very pale through deeper colors to vivid (high chroma) or deep ones (dark colors of high chroma) with a maximum lightness at about 560-mμ center, then become paler again. For the transmission bands, the colors pass from dusky colors (dark and weak) to brilliant colors (light and strong) for moderately wide bands or vivid color (medium

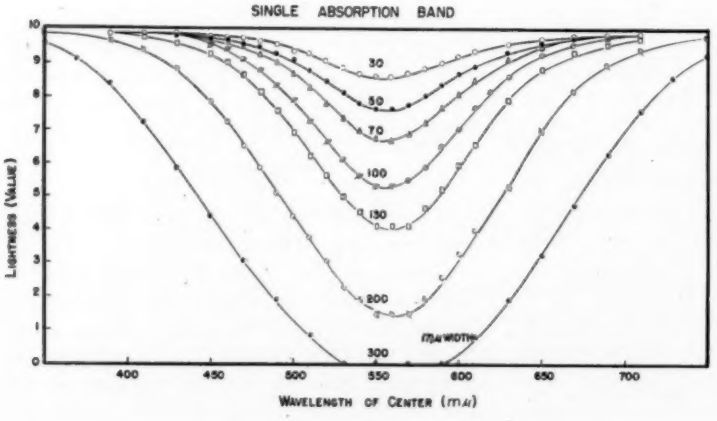

FIG. 7. Munsell value *vs.* wave-length centers of absorption bands for various widths (transmission-band curves are similar but convex upward).

light and very strong) for narrow bands, then darken again and become weaker after passing through a maximum at about 560 mμ.

A true green can be obtained only by means of *two* absorption bands (which may be equivalent to a single transmission band near the center of the visible spectrum).

The maximum possible saturation (chroma) for a given hue and lightness is very nearly reached in the case of some actual present-day yellows, but is very far from being reached in the case of greens and purples. The colors of present-day dyes of other hues fall in between these extremes.

SUMMARY

In order to obtain a more accurate and detailed picture of the effects of changes in absorption spectra on the resulting colors, calculations were made of the colors (for the average person) produced by idealized (rectangular) absorption bands, of various widths, when moved across the spectrum. The calculations are analogous to computations of the maximum possible efficiencies of heat engines calculated from operating temperatures by means of the Carnot cycle.

It has been found that no true greens can possibly be made with a single absorption band,[7] and no magentas with a single transmission band. "Brightness" (saturation) of dyes increases with width of absorption or narrowness of transmission bands; but a law of diminishing returns operates, there being a point beyond which it is unprofitable to go. A 5-mμ shift in the yellow has more effect, for example, on the color than a 50-mμ shift in the blue-green. In two spectral-transmission regions it makes little difference whether very wide or very narrow bands are used; in others the reverse is true. Saturation maxima were found at 3 absorption wave-length positions, and for transmission at 400-450 mμ. Many previous correlations of spectra and "color" are invalidated because of the "crowding" found of some hues into narrow wave-length regions, and great spread in others.

In practice, yellows approach close to the ideal maximum possible brightness, greens are very far away indeed; the spectral conditions for better greens are outlined.

[7] Godlove, J. Opt. Soc. Am. **33**, 351 (1943).

Pittsburgh Conference on Applied Spectroscopy

The eighth Pittsburgh Conference on Applied Spectroscopy, sponsored by the Spectroscopy Society and the University of Pittsburgh, will be held Thursday to Saturday, November 13 to 15, 1947, inclusive, in the Mellon Institute Auditorium. As in the past, sessions on both emission and absorption spectroscopy are planned. Inquiries should be addressed to Dr. Mary E. Warga, University of Pittsburgh, Pittsburgh 13, Pennsylvania.

Practical Methods of Making and Using Multilayer Filters [1]

MARY BANNING
Institute of Optics, University of Rochester, Rochester, New York

(Received February 3, 1947)

The general method of making an evaporated multiple-layer interference filter from two dielectrics is described for any desired wave-length range. The proper thickness of each layer to give maximum reflection in the desired region and at the desired angle of incidence is first computed, and next the visible reflectivity of such a film at normal incidence. This visible reflectivity is used as the gauge to determine the proper thicknesses of the films and the point at which the evaporation of each layer should be stopped. Practical difficulties encountered in making these filters are also discussed with various ways of surmounting them.

Several examples of special filters for particular purposes are reviewed in detail, including infra-red and polarizing filters.

A GREAT deal of material has been published during the last year on methods of computing the spectral reflection and transmission of multilayer films.[2–5] However, little has yet been said about the making of such filters and their practical applications. The present article describes methods of making multilayer filters and easy "rules of thumb" by which the proper film thicknesses may be computed. For lack of sufficient personal experience, no attempt is made to describe metallic-dielectric films such as are used to increase metallic reflection,[6] or to form selective transmission filters,[7] although the same problems occur in these cases. The exact theoretical treatment of both metallic and dielectric multilayer films is given in a forthcoming article by H. D. Polster, whose work coincided with a large part of that described here.

The first highly reflecting films of zinc sulfide were made by Pfund in 1933,[8] and the first "non-reflecting" films of low index fluorides evaporated in 1935 by Strong;[9] in 1939, Cartwright and Turner tried a combination of low and high index films and produced selective reflection.[10] Since that time many laboratories have been working on this problem, but because of the war little has been published except for a short account from the RCA Laboratories.[11]

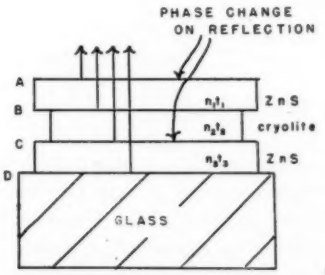

FIG. 1. Schematic diagram of three-layer filter.

SIMPLE FILTERS

Development of multilayer filters at the University of Rochester arose accidentally, in September, 1942, from the need of producing a reflex plate that would reflect approximately 50 percent of the incident light and transmit the rest. A single coating of zinc sulfide, evaporated to an optical thickness of a quarter wave-length for green light, did not give sufficient reflection; if a plate of glass was coated on both sides, the reflection was nearly satisfactory, but this necessitated exposing one of the soft coatings to rough handling and atmospheric conditions. Two plates coated on one side and mounted with their coated faces together were difficult to align per-

[1] The work described here was carried out in whole or in part under contract with Division 16 of the Office of Scientific Research and Development, at the University of Rochester.

[2] P. King and L. B. Lockhart, J. Opt. Soc. Am. 36, 513 (1946).

[3] D. L. Caballero, J. Opt. Soc. Am. **36**, 710A (1946).

[4] R. L. Mooney, J. Opt. Soc. Am. **36**, 256 (1946).

[5] H. D. Polster, J. Opt. Soc. Am. 350A (1946).

[6] A. F. Turner, J. Opt. Soc. Am. **36**, 711A (1946).

[7] Such filters are now being commercially produced by the Farrand Optical Company, Inc., New York City.

[8] A. H. Pfund, J. Opt. Soc. Am. **24**, 99 (1934).

[9] J. Strong, J. Opt. Soc. Am. **26**, 73 (1936).

[10] C. H. Cartwright and A. F. Turner, Phys. Rev. **55**, 1128A (1939).

[11] G. L. Dimmick, J. Soc. Mot. Pic. Eng. **38**, 36 (1942).

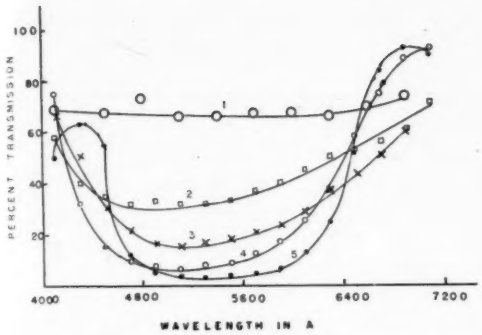

FIG. 2. Effect of addition of layers, first-order interference. (1) 1 ZnS layer; (2) 2 ZnS and 1 cryolite; (3) 3 ZnS and 2 cryolite; (4) 4 ZnS and 3 cryolite; (5) 5 ZnS and 4 cryolite.

fectly parallel and double images resulted, while cementing the two together decreased the reflection too much.

At the suggestion of J. W. Evans, a low index dielectric layer of cryolite was evaporated on top of the high index layer and followed by a second high index film, thus "substituting" for the intermediate air layer. It was soon found that the thickness of the low index layer was as critical as that of the high index layers and that best results were obtained when all three had an optical thickness of a quarter wave-length. Such a film showed 60 percent reflection for the middle portion of the visible spectra.

The explanation of this is very simple. Figure 1 shows a schematic three-layer filter. According to a convention adopted here, high index films are shown with greater length than low index films. Light reflected from all four interfaces, A, B, C, and D, must be in phase for maximum reflection. A phase change of $\lambda/2$ will take place whenever light is reflected from a high index medium into a lower index medium, or at surfaces A and C in Fig. 1. The ray reflected at A suffers a phase change of $\lambda/2$. The ray from B goes $2n_1t_1$ farther than that from A and thus differs in phase by $2n_1t_1$ plus $\lambda/2$; now if n_1t_1 equals $\lambda/4$, the two rays differ by λ and are in phase. The ray from C changes phase by $2n_1t_1$ plus $2n_2t_2$ plus $\lambda/2$ at reflection and is, therefore, $2n_2t_2$ plus $\lambda/2$ out of phase with that from A; to be in phase n_2t_2 must be $\lambda/4$. Similarly n_3t_3 must be $\lambda/4$.

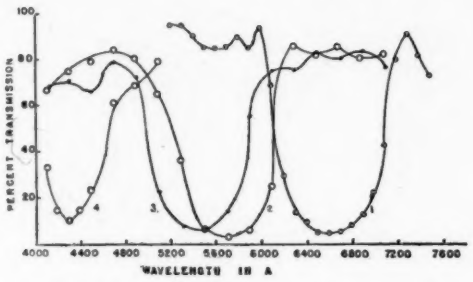

FIG. 3. Second-order filters of seven layers. (1) ZnS to second magenta; (2) ZnS to second yellow; (3) ZnS to second white; (4) ZnS to second blue, first bright blue.

In the case mentioned above, that of a reflex plate using visible light, the wave-length for which this must hold is approximately 5400A. The effect of adding more and more layers is shown in Fig. 2, with the maximum reflection becoming greater and the width of the reflection band becoming narrower. The graphs are shown as transmission values. The reflection is obtained by subtracting these values from 100 percent, as the absorption loss is negligible. This filter is the simplest to make, as the color changes are quite definite and easy to see. It is the fundamental *first-order* filter, with the optical thickness of all layers a quarter wave-length for green light. It is not advisable to put down more than four sulfide layers (with three of cryolite), as the total thickness becomes so great that the film tends to crackle off. This is even more pronounced with layers of greater thickness, as described below.

Reference to Fig. 1 shows that if all the films are made twice as thick ($nt=\lambda/2$), the ray from

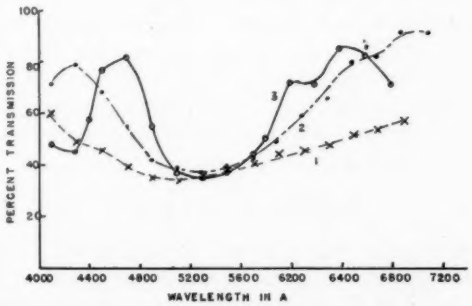

FIG. 4. Effect of changing the order of cryolite in a three-layer filter. ZnS always first order. (1) cryolite first order; (2) cryolite second order; (3) cryolite third order.

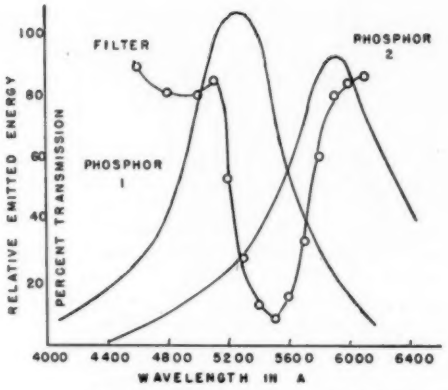

FIG. 5. Multilayer filter designed to distinguish between the emission of two phosphors.

A will tend to cancel that from *B*, while that from *C* will cancel the one from *D*. Thus there will be very little reflection of green light. However, light of twice this wave-length, about 1.1 microns, will find the optical thickness equal to $\frac{1}{4}$ of its wave-length, and will be strongly reflected. This filter has been found useful as a reflex plate for stimulating and observing infra-red phosphors which have a maximum sensitivity to radiation of 1.1 microns, and emit in the green. Now if the films are made even thicker so that $nt=\frac{3}{4}\lambda$, the four rays will again be in phase and again green light will be strongly reflected. This is called a *second-order* filter, several of which (of 7 layers) are shown in Fig. 3, for light of various wave-lengths. It can be seen that the reflection band is much narrower than is the one of a first-order filter. Going to the third order, or $\frac{5}{4}\lambda$-thickness, gives an even sharper band, useful when only a short spectral range is to be considered. In some cases it is desirable to combine cryolite layers of first order with second-order zinc sulfide layers, or *vice versa*. The effect of changing the cryolite order while holding that of the sulfide constant is shown for a three-layer filter in Fig. 4.

DEPOSITION OF LAYERS

Before taking up more complicated filters, the method of deposition should be mentioned. It is not at all necessary to break the vacuum in an evaporator during the making of a filter; although this does no harm to the filter, the total time consumed is much greater than if all layers are deposited in the same vacuum. A small test piece, shielded from evaporating materials except for a small area, can be placed beside the actual sample and rotated slightly after each layer to present a clean surface for observation. The cryolite and zinc sulfide are placed in crucibles side by side on the bottom of the evaporating chamber and the sample centered above them.

If the sample is large, it can be rotated during evaporation about an axis perpendicular to its face, and the crucibles offset from the center to provide an even coating. The proper crucible

TABLE I. Colors of different thickness of zinc sulfide and cryolite films as seen by reflection at normal incidence.

Color		
ZnS	Cryolite	Optical Thickness for Green Light
bluish white	yellow	
white	magenta	$\lambda/4$, first order maximum
yellow	blue	
magenta	white	$\lambda/2$, first order minimum
blue	yellow	
greenish white	magenta	$3\lambda/4$, second order maximum
yellow	blue	
magenta	greenish white	λ, second order minimum
blue	yellow	
green	magenta	$5\lambda/4$, third order maximum

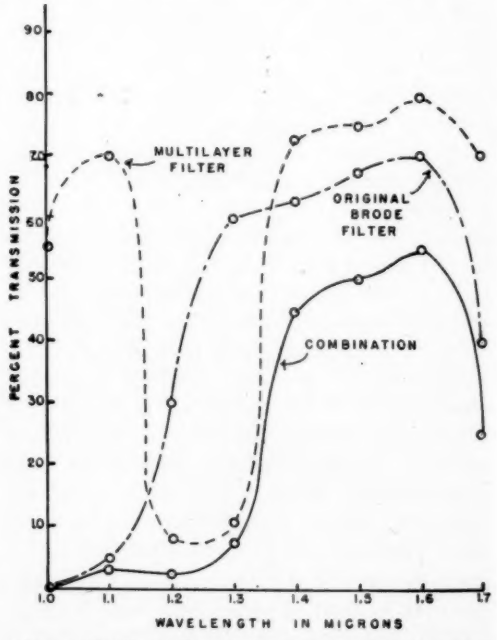

FIG. 6. Multilayer filter designed to sharpen the cut-off of another infra-red filter.

positions are found by successive approximations. Circular filters of 7-inch diameter have been made in this manner, showing only 1 percent variation in reflectivity over their surfaces. This can be done with crucibles only nine inches below the level of the plate, in a 12-inch diameter bell jar. A small hole drilled in the center of the disks allows access to a test piece supported just above the large sample.

The evaporators used here have large flat glass tops for viewing purposes.[12] Thicknesses of the layers are judged by watching, at as near to normal incidence as possible, the changing colors reflected from the sample or test piece. A fluorescent lamp hung over the chamber provided a good white light source. Both the back of the sample and the inside of the glass top are coated with mineral oil to prevent deposition on these surfaces in the form of a reflecting film and consequent confusion in viewing the reflected light.

As the evaporation of zinc sulfide proceeds, it first shows a bluish white reflection, then as it gets thicker it reflects a brilliant white (optical thickness of $\lambda/4$ for green light), yellow, magenta, blue, greenish white, etc. The third "white" is a definite green, as even a single layer shows a much narrower reflection band in the third order. Finally the color changes alternately from pink to green as the orders progress. Cryolite starts with a faint yellow, hard for a novice to see, then turns magenta ($\lambda/4$ for green light), blue, white, deep yellow, deep magenta, blue, green, etc. Table I shows the colors seen at various thicknesses for both the sulfide and fluoride. If a plate is to be used at an angle of incidence other than normal, the colors should be viewed *at that angle*, since at high angles they will differ considerably from those shown at normal incidence.

MORE COMPLICATED FILTERS

If a wide band filter reflecting yellow light is desired, zinc sulfide is deposited to the first yellow reflection, and cryolite to the first blue. A narrower band as shown in Fig. 3 can be made with the second yellow and the second blue. Other visible light filters can be made by following the same law; zinc sulfide to the color desired to be reflected and cryolite to the complementary color.

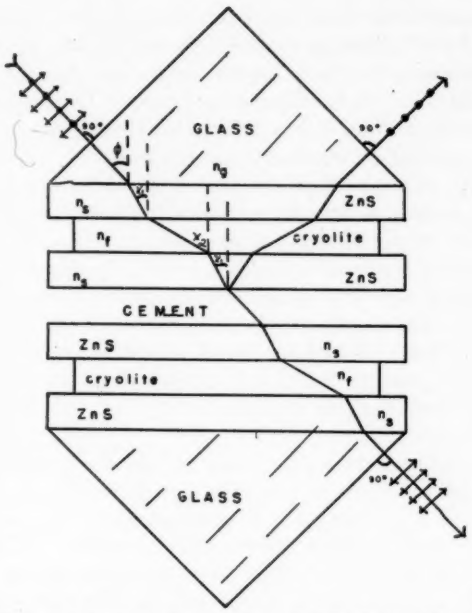

FIG. 7. Schematic diagram of a polarizing beam splitter.

For example, a problem arose involving two different phosphors with separate but close emission peaks. The complete emission bands overlapped to such an extent that distinction between the two was difficult. A filter was desired that would cut out the common region between the two bands and at the same time preserve as much as possible of the rest of the emission. Figure 5 shows the two emission bands and the curve of the filter designed to distinguish them. A green reflection was desired, so white sulfide and magenta cryolite were indicated. Since a very sharp cut-off was needed, it was decided to use third-order cryolite and first-order sulfide. The filter curve shown is the result of the first attempt.

Filters such as these cannot be made for the ultraviolet except for the region very close to the visible spectrum, as zinc sulfide begins to have an appreciable absorption around 3650A. However, the infra-red region is still open to exploration, the only limit imposed is that of the necessary thickness of the layers. Since "colors" in the infra-red cannot be detected easily during

[12] M. Banning and F. W. Paul, Rev. Sci. Inst. **15**, 152 (1944).

evaporation it is necessary to calculate what color of visible light will be reflected at the proper thickness for the infra-red, and watch for this. Although the resulting filter may not be perfect the first time, it is usually close enough to predict what minor changes must be made. A specific example of this occurred when it became desirable to sharpen up the cut-off of another kind of filter, a gel filter made at Ohio State University by Dr. Wallace Brode. As shown in Fig. 6, this was opaque from 0.6 to 1.0 micron and then slowly rose in transmission to 60 percent from 1.3 to 1.6. However, it was to be used for a purpose in which radiation of 1.3 microns was not desired but that of 1.4 microns was necessary.

This called for a second-order filter with maximum reflection at 1.3μ. Since second-order films at 1.3μ are very thick, it was decided to compromise on first-order sulfide and second-order cryolite. To find the wave-length reflected by the sulfide in the visible, the following formula is used:

$$nt = \lambda/4 = 3x/4,$$

where λ is the wave-length in the infra-red and x is the wave-length in the visible. In this case the region around 4300A, or second-order blue, is reflected. Thus the sulfide must be evaporated until it reflects second-order blue.

The cryolite is to be second order in the infra-red; to find what color it will reflect in the visible, the appearance of the sulfide, if evaporated to second order at 1.3 microns, is first calculated. In this case the formula is

$$nt = 3\lambda/4 = 5x/4 = 7y/4 = 9z/4.$$

Supplementary maxima are seen to be at 7800, 5570, and 4300A of third, fourth, and fifth orders, respectively. This would, with sulfide, probably appear to be the fourth blue just before the fourth green. The *complementary* cryolite color is the fourth yellow. The curve obtained with second blue sulfide and fourth yellow cryolite is shown in Fig. 6 along with the resultant curve of the two filters.

Another method for making these high order filters is to compute the color of the sulfide, and then find the cryolite color by trial and error; a good guess is usually approximately right. Once the proper colors are found, they are not at all hard to duplicate. A second person, once shown the colors, can reproduce a filter so that the curves made by two different operators are nearly indistinguishable.

If the filter is to be used at an angle of incidence other than normal, and if for some reason it cannot be watched at this angle during evaporation, a simple calculation gives the color it will reflect at normal incidence. In this case the optical thickness is no longer nt but $nt\cos\theta$, where θ is the angle of refraction in the film. θ_1, the angle of refraction in the sulfide, and θ_2, the angle in the fluoride, are computed from the known angle of incidence and refractive indices by Snell's law. Thus

$$nt = \frac{\lambda}{4\cos\theta} = \frac{\lambda'}{4},$$

where λ is the wave-length reflected at the angle of incidence and λ' is the wave-length reflected normally ($\lambda' > \lambda$).

POLARIZING BEAM SPLITTERS

In January, 1943 we observed marked polarization by reflection from multilayer films on glass plates. In March, 1943 a request was made by Dr. S. M. MacNeille, of Eastman Kodak Company, for the development of polarizing beam splitters made by the deposition of multilayer filters on glass prisms which, as he pointed out, would permit light to enter the film at the polarizing angle. These have been made and found to work quite satisfactorily. Films are deposited on the hypotenuse faces of two 45–90 prisms which are then cemented to form a cube, as shown in Fig. 7. Half of the incident light is reflected and half transmitted, each part being plane polarized to more than 98 percent, over a small angle of incidence.

In order for this to be true, the incident light must strike each interface of the multiple film at the polarizing angle, and the total reflectivity of the film must be such that all of the light polarized perpendicular to the plane of incidence is reflected. It is also desirous that the light path be normal to the entrance and exit faces of the cube.

The index of refraction of the glass, as well as those of the two evaporated substances, is very

important in assuring that the reflection takes place at the critical angle. As shown in Fig. 7, n_g, n_s, and n_f are the indices of refraction of the glass, sulfide, and fluoride, respectively; while φ, x_1 and x_2 are the angles of incidence at the glass-sulfide, the sulfide-fluoride and the fluoride-sulfide interfaces. From Brewster's and Snell's law the following equation must hold for the polarizing angle.

$$n_g^2 = 2n_f^2 n_s^2/(n_f^2+n_s^2).$$

Assuming that zinc sulfide has an index of 2.3 and cryolite of 1.25 (evaporated close to 10^{-3} mm Hg), the index of the glass prisms as computed from this formula must be 1.55. Another factor to be taken into account is the V-number of the glass. Zinc sulfide has a high dispersion with a V-number of 17. Thus in order to make certain that all wave-lengths reach the sulfide-fluoride interfaces at approximately the polarizing angle, the dispersion of the glass must be carefully chosen. It can easily be proved that for the ideal case, assuming the dispersion of cryolite is negligible:

$$V_g = n_g(n_f^2+n_s^2)\,V_s(n_g-1)/n_s(2n_f^2-n_g^2)(n_s-1).$$

Substituting in the equation gives $V_g = 48.5$. Fortunately, either extra light flint or light barium crown will do for the purpose, with the first alternative being by far the best.

Once the best indices of materials and glass (the same process can be repeated for different evaporated materials) are chosen, the proper thickness of the two films remains to be found. A broad reflection maximum, or first-order green, is desired, to cover as much of the visible spectrum as possible. Because total reflection will take place at the hypotenuse faces before the prisms are cemented together, it is impossible to view the prisms at the polarizing angle while evaporating. Therefore, it is desirable to know what color will be reflected at approximately normal incidence if a film of proper thickness is deposited. This is done by computing the angles x_1 and x_2 and using the cosine formula given above. Roughly, the high index film will appear a faint first yellow and the low index film the first greenish white. With extra light flint, zinc sulfide, and cryolite, the ratios of the undesired to desired polarized components are shown in Table II, for both beams and three different colors,

TABLE II. Ratio of undesired to desired components of sample polarizing prism.

	Red	Green	Blue
Reflected beam	.0035	.0065	.0103
Transmitted beam	.0138	.0031	.0545

taken normal to the cube faces. This beam splitter was made with three zinc sulfide layers on each prism. The polarization for white light remains greater than 98 percent over 5 degrees on each side of the normal entrance.

There is no reason why equally good polarizers cannot be made for the infra-red.

Although only zinc sulfide and cryolite are mentioned here, it is solely for the purpose of simplicity. Any high or low index materials will, in general, act the same way. The advantage of sulfide-cryolite films is that they are easy to evaporate and give a large index difference, insuring a large reflection at each interface. However, they are soft and sensitive to atmospheric changes.

The author wishes to express her special gratitude to Dr. J. W. Evans for his constant encouragement and advice on all this work, and to Mr. Robert Hills, Jr., for his assistance during most of the development.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Polarization and Intensity Measurement in Raman Spectra by the Photoelectric Method

D. H. RANK
Physics Department, The Pennsylvania State College, State College, Pennsylvania
(Received May 12, 1947)

A method is described for obtaining the true relative intensities of lines in a Raman spectrum excited by natural unpolarized light. Detection and intensity measurement was accomplished by means of a large photoelectric grating spectrograph. Correction of intensity measurements for the variable polarization and line breadth of the Raman lines is accomplished by methods described in this paper. Depolarization factors have been measured for natural unpolarized exciting light (ρ_n), for perpendicularly polarized exciting light (ρ_s), and for parallel-polarized exciting light (ρ_p). The spectra investigated are those of CCl_4, $CHCl_3$, C_6H_6, and CS_2. The values obtained for ρ_n agree well with those obtained by numerous investigators. The values obtained for ρ_s and ρ_p seem to be in accord with the predictions of the theory. ρ_p seems to be highly sensitive to changes in symmetry of highly symmetrical vibrations, reaching a very large value for vibrations where $\rho_n = .06$ and rapidly approaches 1.00 for vibrations where $\rho_n = .30$ as well as for completely depolarized lines.

Measurements of ρ_n, ρ_s, and ρ_p were made for the Rayleigh scattering of these same molecules, using the identical apparatus employed for the Raman spectrum. Qualitatively the measurements are satisfactory. The presence of parasitic light shows that modifications will have to be made in the scattering apparatus to obtain satisfactory quantitative ρ values, especially for highly isotropic molecules.

PRECISION measurements of depolarization factors are exceedingly difficult when use is made of photographic photometry for the measurements of the intensities of the spectrum lines.[1] Cabannes[2] and Cabannes and Rousset[3] have given the necessary ground work and developed precision methods for the experimental determination of depolarization factors. A discussion of various experimental arrangements for making these measurements may be found in Kohlrausch's book, *Ramanspektren*.[4]

Edsall and Wilson[5] have made depolarization measurements making use of polarized incident light instead of natural unpolarized light, as is ordinarily the custom. The author and his collaborators[6] have demonstrated a totally different experimental method of measuring depolarization factors, which method is capable of yielding precise results. Since the publication of the last mentioned paper, development of the photoelectric Raman spectrograph[7] has been considerably perfected. This instrument has been used continuously for analytical purposes for several years. The resolving power and response of the photoelectric instrument are comparable to the usual photographic technique. The ability to measure intensities directly and precisely, without an intermediate photographic process, is an advantage of no small consequence for polarization measurements. It is the purpose of this paper to measure the relative intensities and the depolarizations for unpolarized and polarized incident light for a number of simple molecules.

The method used for the excitation and recording of the Raman spectra investigated is identical with the full description given in a previous paper.[7] The only modification made in the present investigation was that a stop was placed over the window of the Raman tube for polarization measurements to further limit any parasitic light which might be present. This precaution probably was unnecessary, since tests showed that no detectable difference in depolarizations was measurable with and without the auxiliary stop.

INTENSITY MEASUREMENTS

True relative intensity measurements of the Raman lines would furnish information concern-

[1] J. H. Hibben, *The Raman Effect and its Chemical Applications* (Reinhold Publishing Corporation, New York, 1939), p. 35.
[2] J. Cabannes, Comptes Rendus **187**, 654 (1928).
[3] J. Cabannes and A. Rousset, Ann. de Physique **19**, 229 (1933).
[4] K. W. F. Kohlrausch, *Ramanspektren* (Akademische Verlagsgesellschaft Becker and Erler, Leipzig, 1943), p. 40.
[5] J. T. Edsall and E. B. Wilson, J. Chem. Phys. **6**, 124 (1938).
[6] D. H. Rank, R. J. Pfister, and H. H. Grimm, J. Opt. Soc. Am. **33**, 31 (1943).
[7] D. H. Rank and R. V. Wiegand, J. Opt. Soc. Am. **36**, 325 (1946).

ing the transition probabilities for these molecular energy changes which are observed by means of the spectrum. Any calculations of intensities, of course, do not take into account incorrect experimental procedure, which has been used almost without exception for the measurement of intensities. The chief sources of error (which are large) in the usual intensity measurements, aside from photometric errors, are twofold: (1) The variable polarization of the different Raman lines; (2) the variable natural breadth of the lines.

Considering (1) above it is easy to see that the apparent relative intensity of polarized and depolarized lines will depend on the transmission of the spectrograph for the two kinds of plane-polarized light. This transmission factor will vary widely with different spectrographs, depending on the character and arrangement of the optical components of the dispersing system. In addition, different methods of excitation of the spectra can produce abnormal polarizations of the lines in question. With regard to (2) it can be seen that the apparent intensity of a line, as merely measured from the relative photographic blackening, will, aside from (1), vary with the dispersion of the spectrograph employed in the case of broad lines. Even if the spectrograph and dispersion is kept constant, there will be considerable apparent variations in relative intensities for broad and sharp lines with slit width for a number of optical reasons. In view of the factors mentioned above it is not at all surprising that relative intensities of Raman lines measured in the ordinary way may spread 100 percent more or less.

Correction for the errors enumerated in (1) above was made in the following manner in the present investigation. The method of excitation used was known to approach closely theoretical conditions for the production of spectra exhibiting correct theoretical depolarization factors. Spectrograph and projecting system were calibrated for their transmission of the two kinds of plane-polarized light.

In the following discussion on polarization the conventions used by Kohlrausch, p. 23 of reference 4, will be used with regards to the vibrations of the electric vector.

The transmission of an instrument T for natural unpolarized light is

$$T = S_0/P_0, \tag{1}$$

where S_0 is the intensity measured for the $\perp$ component, i.e., electric vector vibrating in the vertical plane, and P_0 the intensity for the $\parallel$ component, i.e., electric vector vibrating in the horizontal plane. Then ρ_n, the depolarization factor for a polarized line, will be given by

$$\rho_n = \perp/\parallel = S_0/P_0T. \tag{2}$$

If I_t represents the true intensity of the line in question and I_0 represents the observed natural intensity of the spectrum line, the following relationships result.

$$I_t = P_0(1+\rho_n), \tag{3}$$

$$I_0 = P_0(1+\rho_n T), \tag{4}$$

from which is obtained

$$I_t = I_0(1+\rho_n)/(1+\rho_n T).$$

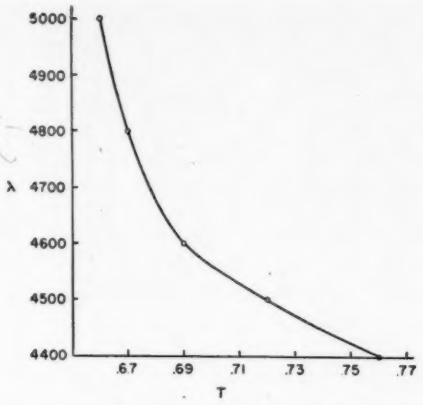

FIG. 1. Transmission curve of the photoelectric spectrograph for plane-polarized light. λ wave-length in angstroms. T is the transmission factor $= S_0/P_0$.

TABLE I. Relative intensities of the Raman lines of CCl_4, $CHCl_3$, C_6H_6, and CS_2. The numbers in columns under molecules represent wave-number shift from the exciting line. Numbers in columns headed I represent relative intensity on a basis of 100 for the strongest line.

CCl_4	I	$CHCl_3$	I	C_6H_6	I	CS_2	I
218	66	262	100	606	12	396	1.3
314	88	366	70	850	4	648	44
459	100	668	80	992	100	656	100
762	37	761	56	1176	18	796	15
790	37	1215	15	1584	13	805	9
1539	5	1505	2	1605	9		
		3019	30	3047	45		
				3063	74		

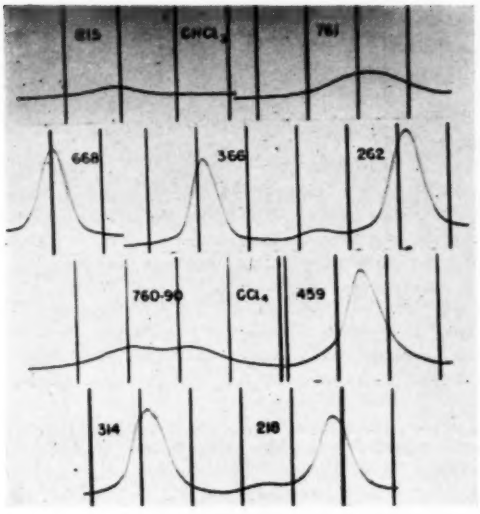

FIG. 2. Reproductions of records from the photoelectric spectrograph showing the stronger lines in the Raman spectrum of CCl_4 and $CHCl_3$. The numbers in the figure refer to wave number shift from the exciting line. Excitation λ 4358Hg.

Spectrograph and projecting system were calibrated for transmission of the two kinds of plane, polarized light in the following manner. A small lamp was operated at constant voltage and placed in the position normally occupied by the scattering tube. The response of the spectrograph to $\perp$ and $\parallel$ polarized light was then measured for a number of wave-lengths. T was found to vary with wave-length from 0.76 for $\lambda = 4400$A to 0.66 for $\lambda = 5000$A. The results of these measurements are given in the curve, Fig. 1.

Normal procedure in photographic microphotometry is to have the width of the image which scans the photographic plate smaller than the breadth of the photographic line images. Such procedure produces "flat-topped" photometer curves and these curves closely approximate the density contour of the lines in question. Occasioned by the small amount of light available to the photoelectric spectrograph, it is necessary to have the exit slit sufficiently wide so that all the light from a monochromatic image will fall between its jaws. Under these conditions the resulting wave-length *vs.* intensity curve will not truly represent the intensity contour of the line in question. However, "peak heights" will be proportional to the intensity of lines which have the same contour. The area under the wave-length *vs.* deflection curve will be a measure of the intensity of the lines regardless of its contour.

Excitation of the Raman spectrum was by λ4358 group of the Hg arc. Wave-lengths shorter than λ4358 were effectively removed by the $NaNO_2$ filter first used by Pfund.[8] It was known from previous work[7] that the intensity of the incident light could be held constant to a few percent for long periods of time. For the purpose of the intensity measurements the spectrum was scanned at the rate of about 1.3A per minute. Results of the intensity measurements corrected for polarization and variation of sensitivity of the photo-tube with wave-length[9] are recorded in Table I. These results are not quite correct in that photo-tube and spectrograph should have been calibrated for wave-length response together. It is felt from the known properties of the grating employed that the results are sub-

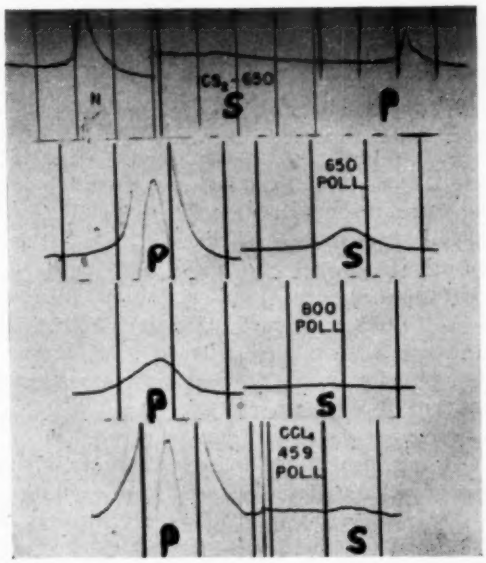

FIG. 3. The upper trace is a reproduction of trace of $\Delta\nu = 650$ of CS_2 as obtained with high resolution. N, no analyzer. S and P refer to analyzed components of the scattered light. The second trace was obtained with normal resolution using S-polarized exciting light. Traces 3 and 4 have a similar origin to that of 2 and refer to the $\Delta\nu = 800$ frequency shift of CS_2 and $\Delta\nu = 459$ frequency shift of CCl_4, respectively.

[8] A. H. Pfund, Phys. Rev. **42**, 581 (1932).
[9] IP21 Multiplier Phototube RCA Victor Division IP21–9–43.

stantially correct to better than 5 percent for lines less than 1600 cm^{-1} shift from the exciting line.

Typical intensity-contour curves are reproduced in Fig. 2. These curves are photographic reproductions of original records as obtained from the photoelectric spectrograph. The heavy vertical lines are wave-length markers and are 5 angstrom units apart.

POLARIZATION MEASUREMENTS

Three sets of polarization measurements were made on the Raman spectrum of CCl_4, $CHCl_3$, C_6H_6, and CS_2. Illumination in the first case (a) was by means of natural unpolarized light, the results of which allow ρ_n to be determined. In case (b) illumination was by means of light polarized so that the electric vector vibrated in a plane $\perp$ to the axis of the scattering tube, which measurement yielded ρ_s. In the third case (c) the light was polarized so that the electric vector vibrated in a plane parallel to the axis of the scattering tube, which measurements yielded ρ_p. Scattering was observed in all cases at right angles to the direction of the incident light.[7] Polarization of the incident light was accomplished by placing a sheet of Polaroid between the cylindrical lens and the scattering tube.

The general theory of light scattering has been treated by Placzek.[10] Summaries of the results of the theory can be found in the books by Kohlrausch[11] and Herzberg.[12] The theory demands for Rayleigh scattering for completely isotropic molecules $\rho_n=0$, $\rho_s=0$. For completely anisotropic molecules $\rho_n^{max}=\frac{1}{2}$, $\rho_s^{max}=\frac{1}{3}$. For modified or Raman scattering for totally symmetrical vibrations $\rho_n=0$ and $\rho_s=0$. Degenerate or unsymmetrical vibrations should yield $\rho_n=6/7$ and $\rho_s=\frac{3}{4}$. There are, of course, intermediate cases between completely isotropic and anisotropic in which intermediate values of ρ are to be expected.

TABLE II. Depolarization measurements with natural and *P*- and *S*-polarized exciting light.

	Natural light			*S*-Polarized light			*P*-polarized light			*S*-pol.
$CCl_4\Delta\nu$	S_0/T	P_0	ρ_n	S_0/T	P_0	ρ_s	S_0/T	P_0	ρ_p	P_0
218	24.6	27.6	.89	6.9	9.9	.70	5.4	5.5	.98	
314	28.1	31.9	.88	8.5	11.9	.71	6.2	6.1	1.02	
459	4.2	65.4	.064	.95	36.1	.026	1.08	.15	7.2	18.0
762	6.5	7.55	.86	1.94	2.55	.76	1.8	1.7	1.06	
790	6.5	7.55	.86	1.94	2.55	.76	1.8	1.7	1.06	
$CHCl_3\Delta\nu$										
262	18.4	21.4	.86	5.7	7.6	.75	5.9	5.9	1.00	
366	4.8	25.0	.19	1.5	14.5	.10	1.6	1.2	1.33	
668	2.4	29.3	.082	.63	17.8	.035	.83	.4	2.07	17.5
761	5.1	5.5	.93	1.5	2.0	.75	1.7	1.5	1.13	1.9
1215	2.0	2.4	.83	.6	.9	.67	.72	.55	1.30	
3018	2.3	9.6	.24	.23	2.0	.12	.68	.65	1.05	
$C_6H_6\Delta\nu$										
606	4.0	4.6	.87	1.4	1.9	.74	1.6	1.5	1.07	
850	1.0	1.0	1.00	.5	.5	1.00	.55	.6	.92	
992	5.1	58.7	.087	2.0	47.6	.042	3.0	1.8	1.67	53.7
1176	3.8	4.6	.83	1.6	2.2	.73	1.8	1.9	.95	
1584	3.1	3.6	.86	1.2	1.5	.80	1.5	1.6	.94	
1605	2.4	2.8	.86	.9	1.1	.82	1.3	1.2	1.08	
3047		dp		dp			4.5	4.2	1.07	
3063	6.8	25.6	.27	2.7	18.5	.15	3.6	3.3	1.09	
$CS_2\Delta\nu$				Unresolved doublets						
650	16.0	57.2	.28	3.15	23.1	.14	10.4	9.0	1.15	68.0
800	2.4	11.0	.22	.42	3.8	.11	1.4	1.4	1.00	13.3
Rayleigh 4916 CCl_4	3.0	28.4	.10	1.7	43.9	.04	*52.2	*16.5	*3.16	
Rayleigh 4916 $CHCl_3$	19.0	63.5	.30	7.1	42.3	.17	9.5	6.6	1.44	
Rayleigh 4916 C_6H_6	18.2	46.6	.39	5.8	29.2	.20	22.0	17.9	1.23	
Rayleigh 4916 CS_2	33.2	62.0	.54	20.6	46.1	.45	46.0	43.3	1.06	

* Indicates measurements were made with λ4358Hg.

[10] G. Placzek, *Marx Handbuch der Radiologie* (1934), Vol. 6, part 2.
[11] See reference 4, p. 25.
[12] G. Herzberg, *Infra Red and Raman Spectra* (D. Van Nostrand Company, Inc., New York, 1945), p. 246.

The above remarks also apply to symmetrical and unsymmetrical vibrations. In addition to the above depolarizations in Raman scattering, it is to be expected from the theory that pure-rotation Raman lines and rotation-vibration Raman lines should be completely depolarized.[13]

The main results of this investigation are summarized in Table II. The data on the molecules investigated are arranged in order of departure of the molecules from being completely isotropic. The galvanometer deflections in cm are recorded in columns headed S_0/T and P_0 so as to furnish an idea of the accuracy of the ρ values. It is believed that the probable experimental error in the deflections will be 1 mm or not over 3 percent, depending on which quantity is the larger. This dual standard of estimating errors is, of course, occasioned by the large variations of relative intensity of lines in a single spectrum. Headings of the first line of Table II refer to the polarization of the incident radiation. The last column in the table is included to show the relative magnitude of the scattering of the two kinds of plane-polarized incident light.

It can be seen from Table II that for completely unsymmetrical vibrations ρ_n and ρ_s approach very closely their theoretical limits. For symmetrical vibrations it can be seen that in no case has the theoretical value for ρ, namely zero, been closely approached, although ρ both for natural incident light and S-polarized incident light is smaller, the greater degree of symmetry of the vibration. There are undoubtedly at least two factors which prevent ρ_n and ρ_s from reaching zero in the present experiments, even in the most symmetrical case, namely CCl_4. (1) The exciting light is not strictly parallel. (2) The isotopic constitution of the various carbon tetrachlorides will remove the over-all symmetry to a slight extent. Finally, it must be remembered that the theoretical values quoted above for ρ_n and ρ_s are derived for the gas. It would be surprising if the liquid state would have a zero effect on the depolarizations, particularly in the case of the symmetrical vibrations.

Probably one of the most interesting features of these experiments with polarized exciting light is the behavior of the Raman spectrum when parallel polarized exciting light is used. Placzek[10] has given the formulae for the ratio of intensities to be expected when the Raman spectra are excited with the two kinds of plane-polarized light. These formulae predict intensity ratios of 6/7 for unsymmetrical vibrations and 0 for symmetrical vibrations. Edsall and Wilson[5] have made qualitative use of polarized exciting light to show the depolarization of the Raman spectrum of a number of molecules. The results observed with P-polarized exciting light can be seen by reference to Table II. ρ_p approaches 1.00 for unsymmetrical vibrations and reaches a large value for highly symmetrical vibrations. The total scattering intensity for highly symmetrical vibrations decreases rapidly with increased symmetry when P-polarized exciting light is used. Although ρ_n and ρ_s do not change to any great extent for the most symmetrical vibration of CCl_4, $CHCl_3$, and C_6H_6, the behavior of ρ_p makes the difference of symmetry at once apparent. It must be remembered that probably the ρ_p values given in the table are minimum values, since any mixture of S-polarized light in the P-polarized exciting light would tend to raise P_0 and thus depress the value of ρ_p. In the case of CCl_4 this above mentioned effect is particularly serious, since slightly more than ½ percent of S-polarized light in the P-polarized exciting light would account for the total intensity of the P_0 component observed with P-polarized exciting light (Fig. 3). Indirectly, the situation referred to above furnishes a satisfying secondary check as to the completeness of the polarization of the exciting light.

CARBON BISULFIDE

Polarization of the individual members of the doublets in the spectrum of CS_2 has been investigated by Langseth, Sorensen, and Nielsen.[14] A description of the origin of these doublets and assignment of the states giving rise to them may be found in Herzberg's book.[15]

It was found possible to measure the depolarization of the individual doublet members with the photoelectric spectrograph by narrowing the slits to ⅓ value used for the other work reported in this paper. The results obtained are given in

[13] Max Born, *Optik* (Julius Springer Verlag, Berlin, 1933), p. 402.

[14] A. Langseth, J. O. Sorensen, and J. R. Nielsen, J. Chem. Phys. 2, 402 (1934).

[15] See reference 12, p. 276.

Table III, along with the measurements of Langseth, Sorensen, and Nielsen.[14]

TABLE III. Depolarization of individual members of CS_2 doublets. ρ_nR. refers to measurements made by the author.

$\Delta\nu$	ρ_nL.S.&N.[13]	ρ_nR.
648.3	.15	.16
656.5	.15	.16
796.0	.18	.21
804.9	.18	.21

The results given in Table III are substantially in agreement but do not agree with the results quoted in Table II for the unresolved doublets. The reason for the disagreement was apparent from measurement of the traces. Results given in both Tables II and III are determined from peak values of the intensity curves. Results obtained with high resolution having a greater degree of purity represent the depolarizations of the centers of the Raman bands. It was found that a very appreciable portion of the intensity-contour curves of these bands, removed from the peak, showed an average depolarization of 0.4. This is evidence that the more widely spaced unresolved rotational branches of these bands are depolarized[12] and makes obvious the reason for the apparent discrepancy shown in Tables II and III. It was found that in the case of $CCl_4 \Delta\nu \sim 459$ cm^{-1} the depolarization factor ρ_n was constant across the intensity contour of the line even under the high resolution used for the CS_2 doublets.

RAYLEIGH SCATTERING

Results obtained from measurements of the Rayleigh scattering and presented in Table II are only presented as tentative values. Inspection of the results seem to show definite trends which seem qualitatively to be consistent with the theory. In the case of Rayleigh scattering no intensity problem exists, since it was necessary to use λ4916 instead of λ4358 so as to avoid operating the photo-tube and electrical circuits at levels approximately 100 times that used for the measurements of the Raman spectrum. However, with Rayleigh scattering one is seriously affected by pernicious parasitic light which arises from numerous causes. The values given for ρ_n by R. Gans[16] for C_6H_6 and CS_2 are 0.45 and 0.64, which are materially higher than the present values given in Table II. However, Gans measurements were made using totally different methods and with white light excitation. The values of ρ_n obtained by Gans for CCl_4 and $CHCl_3$ were 0.045 and 0.24, which are considerably lower than the values obtained in the present work. This last discrepancy could easily be explained on the basis of insufficient removal of parasitic light in the present work.

The present arrangement and design of scattering tube and projection system used in conjunction with the photoelectric spectrograph is eminently satisfactory for efficiently producing light flux for procuring Raman spectra. A few percent of unmodified light in addition to the Rayleigh scattering make no difference when Raman spectra are desired. However, this small amount of parasitic light is fatal to the results when Rayleigh scattering is being studied, especially when the depolarization factor is small.

The intensity of the Rayleigh scattering is roughly two orders of magnitude greater than necessary for accurate determination of ρ. This will allow an ample opportunity for adequately removing the parasitic light. It should be possible to determine ρ for gases as well as liquids.

The obtaining of materially lower ρ_n values than Gans in the case of the other two liquids is hard to explain. This discrepancy may probably depend upon the degree of optical "emptiness" of the liquids employed. The subject of these disagreements is under investigation, and once the cause has been found it will be possible to make precision determinations of the three ρ's with considerable ease. In addition, the variation of ρ with wave-length can be measured if such variation exists.

The author is indebted to Dr. R. V. Wiegand and Mr. A. E. Douglas for assistance in obtaining some of the traces for CS_2.

Note added in proof:

Since the above paper was written the strange behavior of ρ_p has been explained. In collaboration with Mr. A. E. Douglas, the highly polarized lines have been reinvestigated. It has been possible to increase the response of the photoelectric spectrograph by nearly two orders of magnitude which allows the investigation of the moderately small convergence errors present in this apparatus. It is hoped to have our complete paper on the theory of convergence errors and the measurement of small depolarization factors prepared for publication in the near future.

[16] R. Gans, *Handbuch der Experimental Physik* (Akademische Verlagsgesellschaft, Leipzig, 1928), Vol. 19, p. 394.

The Total Emissivity of Various Materials at 100–500°C

B. T. Barnes, W. E. Forsythe, and E. Q. Adams
Lamp Development Laboratory, General Electric Company, Nela Park, Cleveland, Ohio

(Received June 9, 1947)

Total emissivities have been obtained for various metal, carbon, paint-coated, and glass samples. The carbon and glass samples were measured at about 500, 320, and 100°C; the others only at 100°C. The equipment consisted of a thermopile, three diaphragms, and three shutters, all at 0°C, and suitable heaters. For calibration, a special blackbody at 100°C was used. Its principal feature was a polished conical brass mirror, also at 100°C, for reflecting blackbody radiation to the thermopile.

The emissivity data for glass were corrected for radiation transmitted through the sample from the blackened heater surface, and for the temperature gradient in the sample. The resulting emissivity values ranged from 0.82 to 0.91 for two glass samples and from 0.67 to 0.775 for fused quartz. For pressed carbon plates, the range was 0.71 to 0.77; for radiator paints, 0.29 to 0.84; for well polished metals, 0.05 to 0.135.

IN a lamp laboratory, one frequently wishes to know what fraction of the input to a lamp is radiated by the bulb. This fraction could be computed if the temperature and emissivity were known for each point on the bulb. To obtain emissivity data for glass and other materials, the set-up indicated by Fig. 1 was constructed. In this figure, *T* is an Eppley surface thermopile mounted in copper rings and covered with a copper diaphragm. The three shutters *S* and the diaphragm above each shutter were also of copper. The construction is such that the lowest shutter and the thermopile case remain at approximately the same temperature as the brass container. The latter was cooled by immersion in ice water or liquid air.

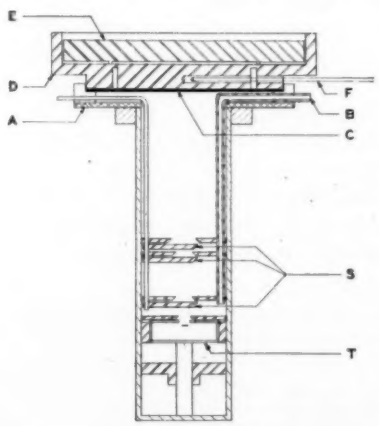

FIG. 1. Set-up for measuring emissivity. *A*, asbestos board; *B*, tube for admitting nitrogen; *C*, sample under test; *D* and *E*, heat-diffusion disks; *F*, thermocouple; *S*, copper shutters; *T*, thermopile.

In order to avoid absorption by water vapor and carbon dioxide in the air, the latter was displaced by admitting dry nitrogen through tube *B*. This was particularly necessary at the lowest sample temperature (100°C) for which data were obtained.

The test sample *C* was made square, for convenience in cutting. It rested on a ledge of brass which was bolted onto the asbestos board *A*. The latter provided thermal insulation between the sample and the thermopile holder. For obtaining sample temperatures greater than 100°C, an electric table stove was placed upside down on a sheet of mica on the copper heat-diffusion disk *E*. The latter rested on a second disk *D*. The air-gap between the two copper disks increases the lateral flow of heat in disk *E*, since it reduces the flow of heat to disk *D*. The circular cut in disk *D* impedes the flow of heat from the central portion of the disk. These two features help make the temperature practically uniform over the central portion of the sample. Since the sample is in direct contact with disk *D* and since the thermal gradient in the latter must be low, the sample temperature is taken to be that indicated by the thermocouple *F*.

For runs at 100°C, the table stove and the heat-diffusion disks were replaced by a thin-walled brass container filled with water and heated by an immersion heater.

The thermopile was connected to a low re-

TABLE I. Total emissivity values for various samples.

Material	Condition	At 100	320	500°C
Alleghany metal	No. 4 polish	.13		
Alleghany alloy No. 66	Polished	.11		
Aluminum	Commercial sheet	.09		
Aluminum	Polished	.095		
Aluminum	Rough polish	.18		
Aluminum paint		.29		
Brass	Polished	.059		
Carbon	Rough plate	.77	.77	.72
Carbon, graphitized	Rough plate	.76	.75	.71
Chromium	Polished	.075		
Copper	Polished	.052		
Copper-nickel	Polished	.059		
Iron	Dark gray surface	.31		
Iron	Roughly polished	.27		
Lamp black	Rough deposit	.84		.78
Molybdenum	Polished	.071		
Nickel	Polished	.072		
Nickel-silver	Polished	.135		
Radiator paint, white		.79		
Radiator paint, cream		.77		
Radiator paint, black		.84		
Radiator paint, bronze		.51		
Silver	Polished	.052		
Stainless steel	Polished	.074		
Steel	Polished	.066		
Tin	Polished	.069		
Tin	Commercial coat.	.084		
Tungsten	Polished coat.	.066		
Zinc	Commercial coat.	.21		

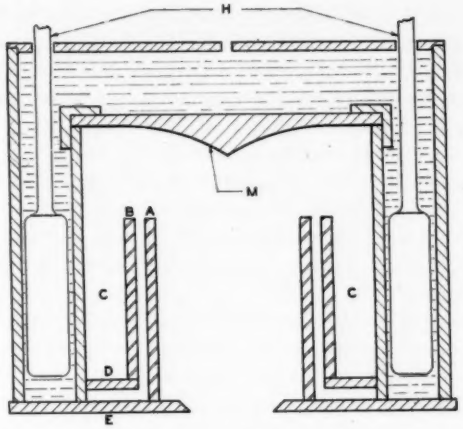

FIG. 2. Blackbody at 100°C. *A*, polished brass cylinder; *B*, brass cylinder blackened on outside, polished on inside; *C*, blackbody cavity; *D*, heavy brass ring; *E*, brass annulus; *H*, immersion heaters; *M*, polished brass mirror.

sistance d'Arsonval galvanometer. The sensitivity of this galvanometer was determined each day. Emissivity values were obtained by comparison of the voltages generated in the thermopile when it was exposed to the test samples with that produced when it was exposed to a blackbody of special design. Figure 2 is a cross-sectional view of this blackbody. It consists of a chamber surrounded, except for the lower side, by a water jacket. The immersion heaters *H* reach almost to the bottom of the water jacket. When the water is boiling vigorously, convection currents in it, and the layer of steam above it, keep the whole mass of water at a temperature of 99° to 100°C.

The copper cylinder *B*, resting on the copper annulus *D*, forms a cavity with walls at a temperature of approximately 100°C. The copper cylinder *A*, which receives heat from the boiler through the base plate *E*, acts as a heat shield for cylinder *B*. Loss of heat from *B* is reduced still further by polishing the surfaces of *A* and *B* which face each other. Radiation from the blackened sidewall of cavity *C* is reflected along the axis of the blackbody by the brass mirror *M*. Since the latter and the source *C* are at the same temperature, the reflectivity of the mirror does not enter into the picture. The aperture of the blackbody was made large enough that the limiting diaphragm was the one under the uppermost shutter, just as in the measurements on the test samples.

For temperatures other than 100°C, the thermopile response with a blackbody source was assumed proportional to $(T^4-T_0^4)$, where T is the absolute temperature of the blackbody and T_0 that of the shutter.

Some runs were made with liquid air around the tube which held the thermopile and the shutter system. In this case, T_0^4 is negligible compared with T^4. However, results obtained with the thermopile at liquid-air temperature were usually less consistent than those obtained

TABLE II. Computation of emissivity values for various glass samples.

Sample	Thickness (mm)	Apparent* emissivity (°C)500	320	100	Computed‡ transmittance 500	320	100	Temperature differential** 500	320	100	Corrected emissivity 475	320	100
Fused quartz	1.96	.78	.80	.75	.266	.134	.023	19	8	1	.67	.76	.775
Corex D	3.40	.80	.80	.76	.113	.041	.002	49	18	2	.91	.90	.83
Nonex	1.57	.82	.82	.78	.145	.041	.004	31	12	1.5	.82	.87	.835

* Assuming no radiation transmitted through sample from heater and no temperature gradient. (Listed temperature is that of heater.)
‡ Assuming all of sample at heater temperature.
** Between front and back surfaces.

with it at 0°C. Further disadvantages in the use of liquid air are the formation of frost at the top of the container and the long period of time required for the thermopile to reach its final temperature.

When the thermopile was at −195°C, its sensitivity was 77 percent of that at 0°C. This was determined by comparison of the ratio of the e.m.f.'s generated in the thermopile at the two temperatures with the corresponding ratio for $(T^4-T_0^4)$. This sensitivity ratio was not needed in the computation of emissivities because blackbody runs were made with the thermopile at each temperature.

The data obtained on various metal, carbon and paint-coated samples are listed in Table I. The carbon samples were pressed plates* with a rough surface. The values in Table I are the respective averages for two kinds of ungraphtized carbon and for two kinds of graphite. In both cases, the difference between corresponding emissivity values for the two samples averaged 0.03.

In Table I, each emissivity value for carbon samples is the mean of several readings. For the other samples only one or two readings were taken. When two determinations were made, the average difference between them was about 0.01 for emissivity values greater than 0.70, and about 0.005 for values less than 0.10. However, systematic errors may amount to as much as 0.03 for values above 0.70, and perhaps 0.01 for the ones below 0.10.

The emissivity values for aluminum listed in Table I refer to three different samples. The highest value (0.18) was obtained on a sample with a rather poor polish. Since even the sample giving the lowest value (0.09) was not smooth, it seems probable that a very smooth surface would give a still lower value.

The sample designated in Table I as tin was commercial tin-plated sheet iron; the "zinc" sample was "galvanized" sheet iron. Some of the other samples were high quality platings on a suitable base metal.

The high value obtained on the nickel-silver sample, as compared with both nickel and copper, may be the result of the zinc content of this alloy.

* We are indebted to Mr. A. C. Downes of the National Carbon Company for supplying these samples.

For the copper-nickel sample, the emissivity value is the average of those for copper and nickel.

Data for 320 and 500°C are not recorded for any of the metal samples because they all tarnished at these temperatures.

THE EMISSIVITIES OF VARIOUS GLASSES

When a partly transparent material such as glass is tested in our set-up, some of the radiation reaching the receiver comes directly from the metal plate which is behind, and in contact with the glass. The total radiation R reaching the receiver is given by the equation:

$$R=R_g+R_mT_g, \tag{1}$$

where R_g is the radiation from the glass, R_m that which would be received from the metal if the glass were not present, and T_g the transmittance of the glass for this radiation.

The transmittance T_g can be calculated if the spectral distribution of R_m is known and spectral-transmittance data for the glass are available. The latter were determined for us by the National Bureau of Standards, for temperatures of 23° and 500°C. The values for 100° and 320°C were obtained by interpolation.

The spectral distribution of R_m was assumed to be the same as that of a blackbody, since the metal was more or less black.** The transmittance values obtained in this manner are listed in Table II.

Substituting these transmittances in Eq. (1), the deflection (R_g) due to the glass alone was found. The deflections obtained with and without the glass samples were substituted for R and R_m in this equation. The ratio of R_g to the deflection which a blackbody at the same temperature would give is the emissivity of the glass.

Emissivity values obtained by the method described in the preceding paragraph may increase with increasing thickness of the glass sample. For wave-lengths for which the absorption coefficient of the glass is very high, all of the radiation will come from the front surface of the glass.

** For the tests at 100°C, the bottom of the brass boiler, directly behind the glass, was coated with a thin layer of soot with an emissivity of 0.84; for the runs at 320 and 500°C, a blackened surface behind the glass was furnished by the oxide layer which formed on the copper-diffusing disk D (Fig. 1). The emissivities of this surface at these two temperatures were 0.65 and 0.575, respectively.

The radiation from the interior of the sample is negligible at 100°C. For wave-lengths for which the sample is partially transparent, radiation will reach the receiver from even the back surface of the sample. As the temperature is increased, the peak of the spectral curve for blackbody radiation approaches the wave-length range in which the glass is partially transparent. At 500°C, a thick piece of glass may have an appreciably higher emissivity than a thin one.

Since glass has a relatively low thermal conductivity, a high rate of transfer of heat through the sample to the radiating surface requires a fairly high temperature gradient. We obtained an approximate value for the temperature difference between the front and back surfaces of our glass samples by means of the following test. First, we measured the radiation from a blackened metal plate at 500°C. Then we put the glass sample between the metal plate and the heater, and again measured the radiation. Now the total radiation is proportional to $e(T^4-T_0^4)$, where e is the emissivity, T is the absolute temperature of the radiation body, and T_0 that of the surroundings. Since e is practically constant over a small temperature range, the reduction in T can be computed from the reduction in thermopile deflection when the glass sample is inserted. When the blackened metal plate is removed, the drop in temperature in the glass plate will be different, unless the emissivity of the glass is identical with that of the black deposit. The ratio of the two temperature gradients will be approximately the same as the ratio of the glass radiation, obtained from Eq. (1), to the radiation from the blackened metal plate with the glass behind it. Temperature differentials between the two surfaces of the glass samples, for a heater temperature of 500°C, are listed in Table II. Since this computation assumes that all of the glass radiation comes from the front surface, the actual temperature differentials will be a little less, and the true emissivities slightly higher than the values given in Table II.

For 320°C and 100°C, the conductivity of the glass was assumed to be the same as at 500°C. (Actually it probably decreases with decreasing temperature.) Then the difference in temperature between the front and the back surface of the glass would be proportional to the radiation R_g from the glass. (See Eq. (1).)

Since most of the radiation comes from the front surface of the glass, we have used the computed temperature of this surface in our calculation of emissivity. Then our data were converted to terms of a common temperature for all of the samples, by use of graphical extrapolation or interpolation. The results are given in the last three columns of Table II.

The emissivity values in Table I increase with increasing temperature when the transmittance of the sample for radiation from a blackbody at the same temperature is less than 10 percent. As the transmittance becomes greater than 10 percent, the emissivity starts to decrease with increasing temperature.

Our data for fused quartz should be roughly comparable with those given by Pirani.[1] The thickness of our sample was practically the same as that of one of his. Our sample probably was somewhat clearer than his, but should have nearly the same emissivity. Table III compares our results. Since his absolute values are computed, we have investigated his method of computation. He takes the emissivity to be 1−transmittance/reflectance. We believe that the proper formula is 1−(transmittance+reflectance). For zero transmittance, the latter formula gives an emissivity value of 0.96, where Pirani's formula gives 1.00. For samples with a transmittance of 10 to 30 percent, the difference is still greater. Consequently, we have taken 0.95 as an approximate correction factor for Pirani's data. For a temperature of 320°C, this gives a result in fairly good agreement with ours. For 500°C, the agreement is practically perfect, in spite of the fact that Pirani's values were computed on the basis of the transmittance at room temperature.

TABLE III. Comparison of our emissivity values for fused quartz with those published by Pirani (reference 1).

	Thickness	Total emissivity At 320°C	At 475°C
Pirani's value†	1.98 mm	.83	.71
0.95×Pirani's* value	1.98 mm	.79	.675
Our value	1.96 mm	.76	.67

† Assuming emissivity = 1 − transmittance/reflectance.
* Assuming emissivity = 1 − (transmittance + reflectance).

[1] M. Pirani, J. Sci. Inst. **16**, 372 (1939).

Power Losses from Heated Objects in Still Air

B. T. BARNES, W. E. FORSYTHE, AND E. Q. ADAMS
Lamp Development Laboratory, General Electric Company, Nela Park, Cleveland, Ohio

(Received June 9, 1947)

The various factors affecting conduction and convection are discussed. Ordinarily, the power carried away from a heated body by these processes cannot be calculated accurately. Two ways of measuring it are described: (1) by boiling water in a container and subtracting from the power input the losses due to evaporation and radiation; and (2) by silvering the bulb of a 500-watt tungsten-filament lamp on the outside, and then making the power input such that the average bulb temperature is the same as that of the unsilvered lamp at normal input. A similar test, but with the bulb silvered on the inside, combined with the results of test (2), gave the power radiated by the bulb. Thus, we find 8 percent of the input to a regular 500-watt lamp carried off by conduction and convection, 11 percent by bulb radiation, the rest by direct filament radiation.

Test (1) showed a conduction and convection loss of about 0.05 watt/cm² at 100°C. This exceeded the radiation loss for every surface tested except a radiator enamel (total emissivity 0.77).

THE heat losses from the surface of a body heated in air or some other gas involve three processes: radiation, conduction, and convection. If the emissivity, the temperature, and the flux density of the back radiation from the surroundings are known, the radiation loss from a flat or a convex surface can be computed quite easily. Losses by conduction and convection ordinarily cannot be calculated because they depend on the shape of a body, on its orientation with respect to the vertical and with respect to surrounding bodies.

In most cases, conduction and convection are successive steps in a single process for removing heat from a surface. A heated body in a gaseous medium is surrounded by a thin layer of gas which is essentially at rest. In air at atmosphere pressure, the thickness of this layer is of the order of one millimeter. Except for radiation, all transfer of heat through this layer takes place by conduction. Outside of it, convection currents help carry the heat away. On horizontal surfaces, since there is little convection, the temperature gradient and, consequently, the heat loss are low. When one uses a strong forced draft, increased convection and, perhaps, turbulence reduce the thickness of the layer of still air next to the heated surface. This increases the temperature gradient and, therefore, the heat loss. The latter is proportional to the product of the thermal conductivity of the gas and the temperature gradient in the layer of still air. Unfortunately, this temperature gradient rarely is known and cannot be measured easily. The most practical way to find the losses due to conduction and convection is to subtract the radiation loss from the heat input required to maintain the heated body at a uniform temperature. In this paper, two sets of experiments which use this basic method are described. These tests were made on cans of square cross section in which water was boiled by means of an immersion heater. These cans were suspended by a string. The power spent in evaporation of water was computed from the loss of weight during a period of 100 minutes; the radiation losses were computed from our emissivity data.* The portion of the input not accounted for by evaporation or radiation must have been carried away by conduction and convection in the surrounding air. Table I shows the results obtained on cans with a surface area of 1500 cm². The average loss by conduction and

TABLE I. The power losses from cans* containing boiling water, in still air at 25°C.

Surface	Emissivity	Input (watts)	Rate of evaporation (g/100 min.)	Radiated power (watts)	Conduction plus convection (watts)	Total surface loss (watts)
Tin plate	.09	180	270	9	74	83
Al paint	.29	200	249	29	82	111
Bronze paint	.51	200	168	50	88	138
Cream paint	.77	200	133	76	75	151

* Dimensions 12.3×12.3×23.7 cm.

* See Table I in the preceding article.

convection was 0.05 watt per cm². The variation in this loss with the type of surface is probably due to the varying degrees of smoothness of the surfaces. The bronze paint had the roughest surface and the unpainted can the smoothest.

The last column of Table I gives the total heat loss from the surface of each can. This varies by a factory of less than two over the range of surfaces shown here, although the emissivity varies by a factor of 8.5. The loss by radiation exceeds that by conduction and convection only in the case of the cream-colored radiator enamel.

Radiator paints of high emissivity, such as the cream enamel which we have tested, make possible the use of less radiating surface than is required with an aluminum paint. At the same time, they may increase the radiant heat to an undesirable degree for a person, or for equipment, near the radiator. Therefore, the choice between radiator paints will depend on whether radiant heat is desired.

POWER LOSSES FOR A 500-WATT LAMP

A regular 500-watt 115-volt tungsten-filament lamp of the type used for general lighting was operated at its normal wattage in still air. The average temperature of its bulb was measured by use of a thermocouple. Then a similar bulb was silvered on the inside with a coat so heavy that no radiation could get through it. A lamp was made in the usual manner except for the use of this bulb. The filament in this lamp was operated at such an input that the average temperature of the bulb was the same as it was in the case of the regular 500-watt lamp. The required input was about 94.5 watts. Except for the loss through the base, all of this power was either radiated by the bulb or carried off by conduction and convection. Since the emissivity of the silver coating is quite low* and since the transmissivity of the bulb for the radiation from the silver is also small, the power radiated from the silver coat through the bulb is negligible.

TABLE II. Distribution of the power output from a regular 500-watt 115-volt tungsten-filament lamp.

	Conduction and convection losses	Bulb radiation	Direct radiation from the filament through the bulb
Percent of input	8	11	81

The outside of the bulb of this silvered "lamp" was then given a good coat of silver. Again, the filament was operated at such an input that the bulb reached the same average temperature as it did in the case of a regular 500-watt lamp. The required input was about 41.5 watts. Subtracting this from the wattage required in the preceding test, we find that the difference between the power radiated by the glass bulb and that by the silver surface is 53 watts. From Tables I and II of the preceding paper, we find that the emissivity of silver is about 0.05 and that of glass about 0.85. Thus the power radiated by the silver surface must be approximately 0.05/0.80 times 53 watts or 3 watts. Then the loss by conduction and convection from a 500-watt lamp must be about 38 watts and the bulb radiation must amount to about 56 watts.

Expressed in terms of the input to the lamp these results are summarized in Table II.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Historical Note on the Rate of a Moving Atomic Clock

HERBERT E. IVES
Bell Telephone Laboratories, New York, New York

(Received May 7, 1947)

The history of the idea of variation of frequency with velocity is followed through Voigt, Larmor, Lorentz, and Einstein. The Michelson-Morley experiment is explainable by any contraction of dimensions in the ratio $(1-v^2/c^2)^{\frac{1}{2}}:1$ along and transverse to the direction of motion. To each contraction corresponds a different value of frequency change. The theoretical speculations pointing to the relation $\nu_m=\nu_0(1-v^2/c^2)^{\frac{1}{2}}$ are discussed, together with the significance of the experimental test by means of canal rays.

THE first suggestion that a "natural" clock would alter its rate on motion appears in a paper by Voigt[1] in 1887. Put in modern terminology his statement translates "We have therefore a glowing plane progressing with velocity c, vibrating with the period $T'=T/(1-v^2/c^2)$." The coefficient of T, it will be observed, is not the same as that proposed by Larmor, Lorentz, and Einstein, and found by experiment,[2] namely $1/(1-v^2/c^2)^{\frac{1}{2}}$. The history and significance of this coefficient and of its exact value, are of some interest.

H. A. Lorentz, following his and Fitzgerald's *ad hoc* postulation of a contraction of material bodies in their direction of motion, to account for the null result of the Michelson-Morley experiment, undertook an extension of the idea "to reduce the equations for a moving system to the ordinary formulae that hold for a system at rest." To do this he introduced new variables:

$$\begin{aligned} x' &= l(x-vt)/(1-v^2/c^2)^{\frac{1}{2}}, \\ y' &= ly, \\ z' &= lz, \\ t' &= l(t-vx/c^2)/(1-v^2/c^2)^{\frac{1}{2}}, \end{aligned} \tag{1}$$

where l is a numerical coefficient, a function of the velocity of translation, whose value is unity for $v=0$. The decision on the value of l to insert was the subject of considerable study by Lorentz, of which more below.

In a note to his *Theory of Electrons* (1913) Lorentz refers as follows[3] to Voigt's publication: "In a paper which to my regret has escaped my notice all these years, Voigt has applied a transformation equivalent to the formulae (287) and (288). The idea of the transformations used above might therefore have been borrowed from Voigt and the proof that it does not alter the form of the equations for the free ether is contained in his paper."

On examining Voigt's original paper it will be found that he did not express his findings in the general form given in (1) but selected for discussion the equations

$$\begin{aligned} x' &= x-vt, \\ y' &= y(1-v^2/c^2)^{\frac{1}{2}}, \\ z' &= z(1-v^2/c^2)^{\frac{1}{2}}, \\ t' &= t-vx/c^2, \end{aligned} \tag{2}$$

that is, he chose the value $(1-v^2/c^2)^{\frac{1}{2}}$ for Lorentz' l. The reason for this choice is not clear.

It has been frequently asserted that the value of l is of no importance. Thus Cunningham[4] states: "The effect of the factor l indicates only a uniform magnification of the scales of space and time, or what is the same thing, a change of units. It does not introduce any essential modification;" and Silberstein:[5] "The value of such a coefficient is essentially, from the physical standpoint, a matter of indifference." While this is so, if the only purpose of the transformations is to explain the Michelson-Morley experiment, *it is not the case if we are to predict the rate of a moving clock.*

This can be made clear by a graphical presentation. Consider the two end mirrors of a

[1] Voigt, "Über das Doppler'sche Princip," Göttinger Nachrichten (March 10, 1887).
[2] Ives and Stillwell, "An experimental study of the rate of a moving atomic clock," J. Opt. Soc. Am. **28**, 215 (1938), and **31**, 369 (1941).
[3] Lorentz, *The Theory of Electrons* (Columbia University Press, New York, 1906), reprint by Stechert, 1923, p. 198.
[4] Cunningham, *The Principle of Relativity* (Cambridge University Press, Teddington, England, 1914), p. 50.
[5] Silberstein, *The Theory of Relativity* (Macmillan Company, London, 1914), p. 119.

Michelson interferometer as lying at the ends of two perpendicular radii of a sphere (Fig. 1) moving in the direction shown, with the velocity v with respect to the medium transmitting the light waves with velocity c. Then, if the radius $=a$, the length of the "vertical" light path, out and back is $2a/(1-v^2/c^2)^{\frac{1}{2}}$, of the "horizontal" path is $2a/(1-v^2/c^2)$, and the times of signal transit are $1/c$ times these. If the sphere is replaced by an ellipsoid with radius in the direction of motion in the ratio $(1-v^2/c^2)^{\frac{1}{2}}$ to the radii at right angles, the signal transit times will be identical, and the Michelson-Morley result is accounted for.

This ratio may be secured by an infinite series of actual dimensions corresponding to different values of l in Eq. (1). Let us investigate several of these.

First take Voigt's choice of l. This corresponds to an oblate spheroid of radius a in the direction of motion, and of radius $a/(1-v^2/c^2)^{\frac{1}{2}}$ at right angles thereto (Fig. 2a). For this the time of signal transit (in both directions) is $2a/c(1-v^2/c^2)$. Using the successive returns of the reflected signal to the origin as clock "ticks" the clock period is

$$T_m = T_s/(1-v^2/c^2)$$

as found by Voigt.

Next, take the value of l finally chosen by Lorentz, namely, $l=1$. This is another oblate spheroid of radius $a(1-v^2/c^2)^{\frac{1}{2}}$ in the direction of motion, of radius a at right angles thereto (Fig. 2b). The clock period is

$$T_m = T_s/(1-v^2/c^2)^{\frac{1}{2}}.$$

As another case take the spheroid of the same volume as the stationary sphere (Fig. 2c), a

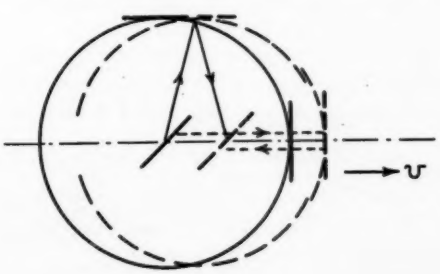

FIG. 1.

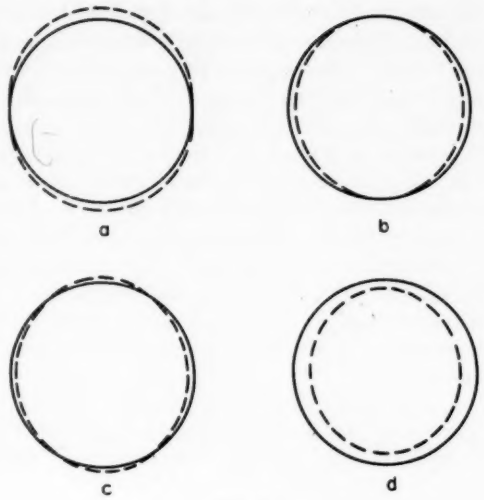

FIG. 2.

form which Lorentz considered attractive for the deformable electron. Here the radius in the direction of motion is $a(1-v^2/c^2)^{\frac{1}{2}}$, at right angles thereto $a/(1-v^2/c^2)^{1/6}$. The value of l is $(1-v^2/c^2)^{1/6}$, and the clock period is

$$T_m = T_s/(1-v^2/c^2)^{\frac{1}{2}}.$$

As a final case consider the spheroid contracted in the direction of motion by the factor $(1-v^2/c^2)$, and by the factor $(1-v^2/c^2)^{\frac{1}{2}}$ at right angles thereto (Fig. 2d). For this the value of l is $(1-v^2/c^2)^{-\frac{1}{2}}$ and the clock period is $T_m = T_s$.

With such a contraction the Michelson-Morley experiment is accounted for, *but there is no change in clock rate.*

For comparison these several theoretical effects are plotted (Fig. 3) in terms of the shift of center of gravity of Doppler lines for approaching and receding canal rays, emitting radiation λ4861, along with the results obtained experimentally.[6] These latter decide unequivocally for $l=1$.

While the matter of the exact value of l has been frequently dismissed as unimportant, it was not so considered by Lorentz, who discussed it at some length. No answer is obtainable

[6] These experimental results are assembled from the two papers given in reference 2, which may be consulted for the units and symbols employed.

through the Michelson-Morley experiment alone.[7] Ultimately Lorentz picked the value $l=1$ by consideration of the variation of mass with velocity as derived from electromagnetic speculations. He was then able to state that his complete system of equations, with $l=1$, "enable us to predict that no experiment made with a terrestrial source of light will ever show us the influence of the earth's motion." Einstein, starting with this conclusion (which was practically Lorentz' working hypothesis from the start), and elevating it to a new principle of physics, was able, by working backward, to deduce the contraction factor $(1-v^2/c^2)^{\frac{1}{2}}$.

The attitude of Einstein toward the consequences of his line of thought was very different in respect to the question of clock rate from that of his predecessors. Voigt gives no reason for developing his transformations. He does not mention the Michelson-Morley experiment. His paper appears to be a mathematical exercise. The conclusion that a moving clock *would* actually behave as is called for by his above quoted statement is not ventured. Lorentz used "local time" "for the sake of facilitating our mode of expression," but never in his original presentations gave to clock behavior the objective reality he maintained for the Fitzgerald-Lorentz contraction. Larmor, who also used local time, came closer when he stated:[8] "The change of time variable in the comparison of radiations in the fixed and moving systems involves the Doppler effect on the wave-length," but left the matter there without further elucidation. Einstein was bold enough to believe what the equations said, and to back up his belief by the prediction that a second-order Doppler effect in canal rays already being sought for by Stark, on rather general grounds,[9] would in fact reveal the relation of frequencies $\nu_m=\nu_0(1-v^2/c^2)^{\frac{1}{2}}$. He entitled his note "Über die Möglickeit einer neuen Prüfung des Relativitätsprinizips,"[10] and accordingly for some thirty years thereafter this observation on canal rays, if and when the experimental difficulties in the way of its achievement could be overcome, figures in discussions of the subject as an *experimentum crucis*.

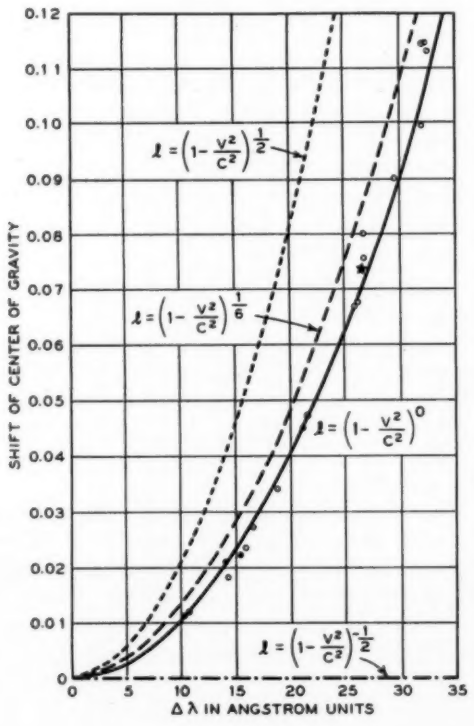

FIG. 3.

[7] Kennedy and Thorndyke, in their experiment with a Michelson-Morley interferometer with arms of unequal length ["Experimental proof of the relativity of time," Phys. Rev. **42**, 400 (1932)] *assume* the linear contraction $(1-v^2/c^2)^{\frac{1}{2}}$ in the direction of motion. Had they assumed the contractions of Fig. 2d they would have "proved" the *non*-relativity of time.

The existence of the motional variation of clock rate, *and its exact value*, are interpretable, from the above account, as decisive verification of the theoretical work of Voigt, Larmor, Lorentz, and Einstein. The whole matter can however, be considered in relation to a new theoretical approach[11] owing nothing to the Michelson-Morley experiment, which latter is the undoubted spark which activated the above investigators.

In this latest approach, radiation is required to conform to Maxwell's equations referred to a wave-transmitting medium, and the requirement

[8] Larmor, *Aether and Matter* (Cambridge University Press, Teddington, England, 1900), p. 177.

[9] Stark, "Über die Lichtemission der Kanalstrahlen in Wasserstoff," Ann. d. Physik **13**, 401 (1906); in particular part III on the probable variation of wave-length according to a function of v^2/c^2.

[10] Einstein, Ann. d. Physik **12**, 197 (1907).

[11] Ives, "Derivation of the Lorentz transformations," Phil. Mag. (7) **36**, 392 (1945).

is imposed that the laws of conservation of energy and momentum shall hold universally and exactly in reactions between radiation and matter. The problem is attacked by study of the mechanical interactions of radiation and matter encountered in the radiation pressure of Maxwell. It is found that the above requirements first demand the variation of mass with velocity, and in turn the variations of dimensions and clock rate according to the factor $(1-v^2/c^2)^{\frac{1}{2}}$ appropriately placed. The impossibility of detecting the earth's motion through the radiation-transmitting medium follows, as in Lorentz' theory, as a consequence. The nul result of the Michelson-Morley experiment appears not as a postulate (as in the special theory of relativity) but as a corollary of older physical laws of wide generality, demanding no "new principle."[12] In the same list of corollaries are the variation of mass and clock rate.

Of experimental tests for establishing the reality of these contractions, however arrived at, the most crucial is the variation of clock rate; for while the Michelson-Morley experiment has the uncertainty that it might be explainable by an entrained ether (Stokes), and the electrical experiments on variation of mass might be explained by variation of charge with velocity, as is recurrently proposed, no comparable alternatives appear to weaken the conclusiveness of the experiment on clock rate.

[12] Newton, *Rules of Reasoning in Philosophy*, "We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearance," Rule 1.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Correspondence Principle Analysis of Double-Photon Emission

JOHN A. WHEELER
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received July 21, 1947)

The process of simultaneous emission of two photons (of continuous probability distribution in energy) in the transition from one discrete quantum level to another is here analyzed in the correspondence principle limit in which the quantum numbers of the two states are large and the change in quantum number is small by comparison. In this limit, it proves possible to define a classical dynamical quantity [Eq. (19)], the frequency analysis of whose classical variation with time in a definite orbit determines in the given approximation the absolute transition probability and intensity distribution in the continuous spectrum in question.

INTRODUCTION AND SUMMARY

A HYDROGEN atom which arrives at the metastable $2s$-state by recombination or otherwise, will be unable to drop down to the fundamental $1s$-state by emission of a single photon of any polarity whatsoever, provided that the atom is located in a region of interstellar space where it is free from the perturbing electric fields of nearby ions and from the exciting influence of a flux of radiation. Under these circumstances the system will nevertheless get to the ground state in a mean time of the order of 1/7 second, according to Breit and Teller,[1] through a process in which two photons of complementary energy come off simultaneously:[2]

$$\hbar\omega_1+\hbar\omega_2=E(2s)-E(1s). \qquad (1)$$

The condition of energy conservation leaves arbitrary the division of the frequency between the pair of quanta. The authors in question calculate the intensity distribution in the continuous spectrum, finding a broad maximum centered at the point of equal photon energies.

The results of Breit and Teller were obtained

[1] G. Breit and E. Teller, Astrophys. J. 91, 215 (1940). *Note added in proof:* Eq. (2) predicts a transition probability smaller by a factor 2π than that given by Breit and Teller. To date the source of the discrepancy has not been located.

[2] Here and below we use ω to designate circular frequency, in units radians per sec.; and use $\hbar$ to represent the quantum of angular momentum (1.027×10^{-27} g cm^2/sec.) as distinct from the quantum of action, h.

by a standard application of the second-order perturbation theory of quantum mechanics in which—to use a purely mathematical manner of speaking—account is taken of the possibility of virtual transitions from the $2s$-state to all p-states, and from these states down to the $1s$-level. The result of the calculation takes a simple form. It yields the probability per second, $dA^{\kappa\lambda}$, of a double-photon transition from the excited state (labeled b) to the ground state (designated a) in which the photon of smaller circular frequency, ω_1, comes off in the interval ω, to $\omega_1+d\omega$, and the one of higher frequency comes off in the complementary energy interval, and in which furthermore (to be more detailed) the "virtual oscillator" responsible for the first photon is polarized along the κ axis ($\kappa=x$, y, or z) and the other virtual oscillator is polarized along the λ axis ($\lambda=x$, y, or z):

$$dA^{\kappa\lambda}=(4e^4/9\pi^2c^6)\omega_1^3\omega_2^3 d\omega/M_{ba}{}^{\kappa\lambda}(\omega)/^2. \quad (2)$$

Here the composite matrix element for the jump $b\rightarrow a$ is given by the sum over all virtual states, v, (of energy E_v) for which the relevant elementary matrix elements do not vanish:

$$M_{ba}{}^{\kappa\lambda}(\omega)=\sum_v x_{bv}{}^{\lambda}x_{va}{}^{\kappa}/(E_v-h\omega_1-E_a) + \sum_v x_{bv}{}^{\kappa}x_{va}{}^{\lambda}/(E_v-h\omega_2-E_a). \quad (3)$$

Here x^{κ} and x^{λ} are the components of the displacement of the radiating electron along the κ and λ axes. When the polarization is not a matter of concern, the total probability per second, dA, of transition with emission into da is found by summing $dA^{\kappa\lambda}$ over all three directions of polarization, κ, and independently over all three directions of polarization, λ. From Eq. (3), and its natural generalization to many-electron atoms, follow certain obvious selection rules for two-quantum emission, such as the requirement that the angular momenta of initial and final states differ either by 0 or by $2\hbar$, and that both states have the same parity.

Apart from these simple deductions from the quantum formulae above, it is difficult to form for a general atomic or molecular system in a particular level of excitation a qualitative idea whether double photon emission is favored or nearly forbidden (as double-photon process goes!).

To translate this question of rate of the two-quantum process from the realm of quantum mechanics to that of classical mechanics, via the correspondence principle, is the purpose of this note. It is well known that the quantum treatment of single-photon emission has a close relation to the classical radiation theory. Likewise, one has to expect that the above formulae for rate of double-photon emission can be set into correspondence with a formula which refers only to quantities which are well defined for a single classical orbit, in such a way that the classical and quantum expressions will come into asymptotic agreement with each other in the limit of high quantum numbers.[3] Specifically, it must be possible, in principle, to find a classical dynamic quantity—in the same sense that coordinates and momenta are such quantities—the frequency analysis of whose classical variation in time for a given orbit will determine the probability for two-quantum emission with an accuracy which is the higher, the closer one is to the correspondence principle limit of large quantum numbers.

We are successful in our goal to the extent that we are able to define a classical quantity of motion which determines the emission probability in the given approximation. We are unsuccessful in the present note in the larger aim of readily visualizing double-photon radiation in the sense that the classical dynamical quantity in question has a not easily grasped relation to such usual features of the motion as displacement and velocity. However, a little insight into the features of the motion responsible for two-quantum emission is acquired by applying the general correspondence principle treatment to two idealized cases of one-dimensional motion (although these particular results could be derived directly from the quantum formulae (2) and (3)). In the case of the harmonic oscillator, double-photon emission is found to have zero probability for all choices of initial and final

[3] For a discussion of the correspondence principle in the light of wave mechanics see, for example, H. A. Kramers, *Quantentheorie des Elektrons und der Strahlung* (Akademische verlagsgesellschaft, Leipzig, 1938), Vol. II, Section 85, also appears in Vol. I, Part II of *Hand- und Jahrbuch der Chemischen Physik*. For a correspondence principle treatment of double-photon processes before the period of wave mechanics, see H. A. Kramers and W. Heisenberg, Zeits. f. Physik **31**, 681 (1925). For a wave-mechanical analysis, see M. Goeppert Mayer, Ann. d. Physik **9**, 237 (1931).

state. In the case of an otherwise free particle which moves back and forth between two rigid walls, the probability of simultaneous emission of two quanta in a continuous spectrum is found in the correspondence principle approximation to be independent of separation of the walls, provided the frequency of collisions with the walls is fixed; and this probability vanishes exactly unless the quantum numbers of the upper and lower levels differ by an even integer. More generally, in the case of a one-dimensional motion, only when the potential is asymmetric with respect to a reversal of the sign of displacements will double-photon jumps be allowed in which the quantum number changes by a single unit.

The remainder of this note is devoted to the derivation of the correspondence principle intensity formula and the examination of the two examples just mentioned.

DETAILS

The starting point for a classical treatment of the radiation from a dynamical system in a given state of motion is usually taken to be a Fourier analysis of the intensity-determining quantities of the motion in terms of the fundamental frequency of the motion, ω, and its higher harmonics, $k\omega$, thus:

$$\text{displacement} = x(t) = \sum_{k=-\infty}^{\infty} X_k \exp(ik\omega t), \quad (4)$$

(where the reality of $x(t)$ requires that $X_{-k} = X_k^*$) and

$$\text{acceleration} = \sum_{k=-\infty}^{\infty} -k^2\omega^2 X_k \exp(ik\omega t). \quad (5)$$

For instance, to take a well-known example, the classical expression for rate of loss of energy by radiation,

$$(2e^2/3c^3)(\text{acceleration})^2_{\text{average}}, \quad (6)$$

takes in the Fourier representation the form

$$\sum_{k=1}^{\infty} (4e^2/3c^3)(k\omega)^4 |X_k|^2, \quad (7)$$

from which follows, in the approximation of the correspondence principle, the probability per second, A_k, of a jump from the given state to a state lower by k quantum units:

$$A_k = (4e^2/3hc^3)(k\omega)^3 |X_k|^2, \quad (8)$$

in direct relation to the accurate non-relativistic quantum formula for dipole radiation:

$$A_{b,b-k} = (4e^2/3hc^3)(W_{b,b-k})^3 |X_{b,b-k}|^2. \quad (9)$$

As this example recalls, we translate from classical theory to quantum theory and conversely by the approximate correspondences:

$$\text{for frequencies:} \quad k\omega \leftrightarrow \omega_{b,b-k}, \quad (10a)$$

$$\text{for energies:} \quad kh\omega \leftrightarrow E_b - E_{b-k}, \quad (10b)$$

$$\text{for displacements:} \quad X_k \leftrightarrow x_{b,b-k},$$

or,

$$T^{-1}\int_0^T dt \exp(-ik\omega t)x(t) \leftrightarrow \int \psi_b^* x \psi_{b-k} dx. \quad (10c)$$

Here T is the period $2\pi/\omega$ associated with the fundamental frequency ω. For simplicity we shall here and hereafter idealize in one index the multiplicity of quantum numbers which will be required in general to specify the typical quantum state, just as in the classical treatment we shall typify in a simple Fourier series, with a single action and angle variable, the several sets of such conjugate quantities which will be needed in actuality to describe the motion. In Eqs. (10) it is not stated whether the fundamental frequency ω and the Fourier components X_k are to be evaluated for that classical motion which possesses bh units of action, or for some intermediate classical orbit. This ambiguity is, of course, typical of the correspondence principle, and the associated uncertainty in the transcription can evidently only then be assumed to be negligible, when the quantum numbers b and $b-k$ are very large in comparison to the quantum change, k.

Proceeding according to the correspondence of Eq. (10), just as in the case of single-quantum radiation, we transcribe the matrix element of Eq. (3) for double-photon emission into classical terms, obtaining

$$\begin{aligned} M_{ba}{}^{\kappa\lambda}(\omega) \rightarrow\ & \hbar^{-1}\textstyle\sum_v X_{b-v}{}^{\lambda} X_{v-a}{}^{\kappa}/[(v-a)\omega - w_1] \\ & + \hbar^{-1}\textstyle\sum_v X_{b-v}{}^{\kappa} X_{v-a}{}^{\lambda}/[(v-a)\omega - w_2] \\ =\ & \hbar^{-1}\textstyle\sum_k X_{b-a-k}{}^{\lambda} X_k{}^{\kappa}/(k\omega - w_1) \\ & - \hbar^{-1}\textstyle\sum_k X_k{}^{\kappa} X_{b-a-k}{}^{\lambda}/(k\omega - w_1) = 0. \end{aligned} \quad (11)$$

This unexpected and zero result shows that we have to go to a higher order of approximation in the application of the correspondence principle to obtain the desired classical formula, and suggests that this formula when obtained will have a less direct connection with the quantum treatment than was the case for single-photon emission.

We would have obtained in a much more familiar part of quantum theory a result likewise equal to zero and likewise misleading if we had applied the correspondences of Eq. (10) without further consideration. We should have found in this way for the commutator of Heisenberg,

$$i\hbar^{-1}\sum_v (p_{mv}x_{vm}-x_{mv}p_{vm}), \tag{12}$$

a value identically equal to zero, instead of the correct figure of unity. Both in this example and in the present problem the means of correction is simple. We need only insert for a quantity such as p_{mv} the $m-v$ Fourier element of the classical quantity $p(t)$ in the mth quantum state, and for x_{vm} the $v-m$ Fourier element of $x(t)$ in the vth quantum state. We need not consider further the example of the Heisenberg commutation relations, where the appropriate analysis has already been given elsewhere,[4] and shall only recall the following general points. First, the Fourier series for the classical variables give their expression in terms of conjugate action and angle variables, $J=(2\pi)^{-1}\int_{\text{cycle}}p(t)dx(t)$ and $w=\omega t$, in the form

$$x(E, t)=\sum_k X_k(J)\exp(ikw). \tag{13}$$

Second, in the equation connecting the angle variable, w, with t, the circular frequency, $\omega=\omega(J)$, is given by the derivative of the energy with respect to the action variable, J. Third, the energy of the kth quantum state is to be considered as given by setting $J=kh$ in the equation $E=E(J)$. Fourth, owing to the invariance of the Hamiltonian equations with respect to change of conjugate variables, we may write the Poisson bracket of any two dynamical quantities, B and C, in any one of the following forms:

$$\begin{aligned}\{B, C\}&=(\partial B/\partial p)(\partial C/\partial x)-(\partial C/\partial p)(\partial B/\partial x)\\&=(\partial B/\partial E)(\partial C/\partial t)-(\partial C/\partial E)(\partial B/\partial t)\\&=(\partial B/\partial J)(\partial C/\partial w)-(\partial C/\partial J)(\partial B/\partial w).\end{aligned} \tag{14}$$

Proceeding in accordance with these principles, we now replace the terms in the matrix element of Eq. (3) as follows, using as our standard of reference the quantities of motion in the bth quantum state:

$$\begin{aligned}&x_{bv}\rightarrow X_{b-v},\\&x_{va}\rightarrow X_{v-a}+(v-b)h(dX_{v-a}/dJ)\\&E_v\rightarrow E_b+(v-b)h\omega+\tfrac{1}{2}(v-b)^2\hbar^2(d\omega/dJ).\end{aligned} \tag{15}$$

Then we find for the desired matrix element the expression

$$\begin{aligned}M_{ba}{}^{\kappa\lambda}\rightarrow&[k\omega-\omega_1]^{-2}[(b-a-k)X_{b-a-k}{}^{\lambda}kX_k{}^{\kappa}\\&\times(d\omega/dJ)]+[k\omega-\omega_1]^{-1}[(dX_{b-a-k}/dJ)\\&\times kX_k{}^{\kappa}-(dX_k{}^{\kappa}/dJ)(b-a-k)X_{b-a-k}{}^{\lambda}].\end{aligned} \tag{16}$$

To obtain an expression for M which has more evident meaning than (16), though no longer its manifest symmetry between the two directions of polarization, we define for each component of displacement, $x^{\lambda}(t)$, of the radiating particle a new associated classical dynamical variable, $s^{\lambda}(t)$, through the equation

$$\dot{s}^{\lambda}+i\omega_1 s^{\lambda}=x, \tag{17}$$

and the supplementary condition that s be periodic in time with the same period, T, which appertains to the displacement, x. Furthermore, we use the Poisson bracket notation of (14), and employ the subscript $b-a$ in connection with any dynamical quantity, $Q(t)$, to denote the $(b-a)$th Fourier coefficient of that dynamical variable:

$$Q_{b-a}=T^{-1}\int_0^T dt\exp(-i(b-a)\omega t)Q(t). \tag{18}$$

Then we have for the desired correspondence principle transcription of the matrix element of Eq. (3) the very simple formula

$$M_{ba}{}^{\kappa\lambda}\rightarrow\{x^{\lambda}, s^{\kappa}\}_{b-a}. \tag{19}$$

[4] See, for example, Ruark and Urey, *Atoms, Molecules and Quanta* (McGraw-Hill Book Company, Inc., New York, 1930), p. 585.

With this result we have achieved our aim of expressing in terms of the quantities characteristic of a single classical orbit the probability for double-photon emission. It is remarkable that the final formula [the result of Eq. (19) inserted into Eq. (2)] contains no mention of the quantum of action, in contrast to the expression (8) for the probability of single-photon emission. As already noted, no easily interpreted visualization of (19) is evident.

Application of our general result to the case of the one-dimensional harmonic oscillator is straightforward. The displacement, $x(t)$, varies with time in a pure sinusoidal manner, with no overtones. Accordingly, the associated dynamical quantity $s(t)$ of Eq. (17) will likewise contain in its Fourier analysis only the two terms $\sin wt$ and $\cos wt$. Consequently s will be representable as a linear combination of the displacement, $x(t)$, and the momentum, $p(t)$. The x part of s will drop out when we evaluate the Poisson bracket $\{x, s\}$, and the p part will give a constant in virtue of the relation $\{x, p\}=-1$. Thus $\{x, s\}$ is equal to a constant. But all the coefficients in the Fourier analysis of a constant are zero, with the exception of the constant term, in which we are not interested. Consequently double-photon emission is forbidden for all transitions in the case of a harmonic oscillator.

In the case of any one-dimensional potential symmetric with respect to the interchange of x and $-x$, the Fourier series for the displacement $x(t)$ will lack all even harmonics. As a consequence, the dynamical quantity, $s(t)$, whatever be the value of ω_1, will also contain odd harmonics exclusively. The Poisson bracket, $\{x, s\}$, being biquadratic in the two quantities, will therefore be built up entirely of even harmonics. Therefore double-photon emission will be forbidden between two states unless their quantum numbers, b and a, differ by an even integer.

In the special case where the potential consists of an infinite wall at $x=\pm L$, and vanishes between these limits, the calculations may be carried through explicitly. In terms of the period, T, of the motion, we take

$$x=(4L/T)(T/4-t) \text{ for } 0<t<T/2,$$
$$x=(4L/T)(t-3T/4) \text{ for } T/2<t<T. \quad (20)$$

Writing θ as an abbreviation for $\omega_1 T$ we find

$$s-ix/\omega_1=(4L/\omega_1{}^2T)\{-1+i[\sin(\theta/2)]^{-1} \times[-1+\exp(-i\theta/2)]\exp(i\omega_1 t)\}, \quad (21)$$

in the interval $0<t<T/2$, and in the next half-period a similar expression, obtained by replacing the first -1 by $+1$, and the second -1 by $-\exp(-i\theta)$.

Whilst $x(t)$ and $s(t)$ repeat themselves only after the interval T, and are continuous functions, the Poisson bracket $\{x, s\}$ is discontinuous at 0, $T/2$, T, etc., and repeats itself after the interval $T/2$. In terms of the time $\tau=t-T/4$, $t-3T/4$, etc., measured from the center of any such interval, we have

$$(mw_1/16T)\{x, s\}=-\theta^{-1}+\exp(i\omega_1\tau) \times\{[2\sin(\theta/4)/\theta\sin(\theta/2)] -[\sin^3(\theta/4)/\sin^2(\theta/2)] -(i\tau/T)[2\sin(\theta/4)/\sin(\theta/2)]\}. \quad (22)$$

The coefficients in the Fourier expression of (22) are readily found, and yield for the probability per second of double-photon emission in the jump from a given level to a state two quantum units lower the result

$$dA\sim(16384/9\pi^2)(e^4/m^2c^6)\omega_1\omega_2 d\omega \times[\sin^4(\theta/4)/\sin^4(\theta/2)]\{1+2\sin(\theta/2) \times[(4\pi-\theta)^{-1}-\theta^{-1}]\}.^2 \quad (23)$$

Comparison of this result with an order-of-magnitude lower limit estimate of the probability per second for single-photon emission,

$$A\gtrsim(e^2/mc^3)\omega^2,$$

shows the very low probability of double-quantum emission. It will be noted, however, that the probability given by (23) diverges when, keeping $\omega_1+\omega_2=2\omega$, we let ω_1 approach the fundamental frequency ω. Obviously the concept of double-quantum emission breaks down when one approaches the values of frequency appropriate to the normal emission of two successive quanta.

On the Intensity of the Light Reflected from or Transmitted through a Pile of Plates

L. B. TUCKERMAN
National Bureau of Standards, Washington, D. C.

(Received July 14, 1947)

A review of the history of the theory of the reflection of light from a pile of plates is given. The presence of incorrect formulae in some current text and reference books is noted. The formulae for a pile of absorbing plates are derived by a method different from that used by Stokes, giving the results in a form different from his. The equivalence of the two forms is proven. An upper bound to the error caused by assuming an infinite instead of a finite number of multiple reflections is derived in the Appendix.

SOME time ago my friend Dr. A. Q. Tool asked me whether I remembered the formulae for the fraction of the intensity of the incident light reflected from or transmitted through a pile of plates such as are used in plate polarimeters. In Drude[1] and Wood[2] he found incorrect formulae which ignored the multiple reflections taking place between the surfaces of the plates.

I did not remember the formulae and rather than spend time hunting down references, I derived them afresh, not merely for wholly transparent, but also for absorbing plates. For transparent plates the solution was simple, involving merely the summation of a geometric series, iteration and induction from n to $(n+1)$ surfaces. It led to formulae, which I recognized as familiar:

$$r_n=n\rho/[1+(n-1)\rho],$$
$$t_n=(1-\rho)/[1+(n-1)\rho]. \tag{1}$$

Here r_n is the fraction reflected from, and t_n the fraction transmitted through n surfaces, and ρ, the reflection coefficient from a single surface, either $\rho_{\parallel}$ or $\rho_{\perp}$.

The method I used in deriving the formulae for absorbing plates was not so simple. It aroused no memory of my having seen it before and the final results seemed wholly unfamiliar. I wondered if they were new and made a search through the literature readily accessible to me. The history which I found seemed to me interesting enough to be worth summarizing. In 1821, Fresnel[3] published his well-known formulae for the reflection coefficients $\rho_{\parallel}$ and $\rho_{\perp}$ of the surface between two media,

$$\rho_{\parallel}=\frac{\sin^2(i-i')}{\sin^2(i+i')}, \quad \rho_{\perp}=\frac{\tan^2(i-i')}{\tan^2(i+i')}, \tag{2}$$

where i = angle of incidence and i' = angle of refraction. To check these, he used measurements of intensity and of the change in the plane of polarization produced by reflection from glass plates. He derived formulae equivalent to Eqs. (1) (reference 3, p. 645) for the effect of the multiple reflections between the two surfaces of a single transparent plate and indicated how they could be extended to more than one plate. In no later publication have I found any reference to this part of his 1821 memoir.

I found a number of later publications in which formulae equivalent to Eqs. (1) were derived. Clausius,[4] in 1848, developed them for a single plate, and Provostaye and Desains,[5] in 1850, for a pile of plates, both without reference to any previous work.

Wild[6] and Bohn,[7] in giving the derivation, credit the formulae to Neumann's lectures in Königsberg.

[1] P. Drude, *Lehrbuch der Optik* (Leipzig, 1912), third edition, pp. 271–272.
[2] R. W. Wood, *Physical Optics* (New York, 1934), third edition, pp. 340–343.
[3] A. J. Fresnel, *Oevres Complètes* (Paris, 1866), Vol. 10, pp. 640–648.
[4] R. Clausius, "Ueber die Intensität des durch die Atmosphäre reflecterte Sonnellichts," Crelle's J. **36**, 185–215 (1848).
[5] F. de la Provostaye et P. Desains, "Memoire sur la Polarisation de la Chaleur par Refraction simple," Ann. de Chim. et Phys. [3] **30**, 158–178 (1850).
[6] H. Wild, "Ueber ein neues Photometer und Polarimeter nebst einigen damit angestellten Beobachtungen," Pogg. Ann. **99**, 235–274 (1856).
[7] C. Bohn, "Zur Polarization des Lichtes durch einfache Brechung," Pogg. Ann. **117**, pp. 117–131 (1862).

Neumann went to Königsberg in 1826 and published his classic memoir on the reflection and refraction of light in 1835, but I have found no early publication of his derivation of these formulae. However, a full discussion is given in his lectures[8] published in 1885.

In 1862 Stokes,[9] also giving no reference to previous work, derived the formulae for a pile of absorbing plates, obtaining formulae equivalent to Eqs. (1) for transparent plates as a limiting case.

Despite all these references to the correct formulae for transparent plates, which I found in my cursory search through the literature, the same incorrect formulae which Tool found in Drude and Wood are found in a number of published papers and current texts and reference books.

Wüllner's *Experimental Physik*[10] although it refers to Provostaye and Desains,[5] gives the same wrong formulae. Although Drude gave the correct formulae in his contribution to Winkelmann's *Handbuch der Physik*,[11] the incorrect formulae given in the first edition (1900) of his *Lehrbuch der Optik* were not corrected in later editions.

Wood, in the 1911 edition of his *Physical Optics*, although in discussing absorbing plates he copies tables and quotes extensively from Stokes' correct treatment (pp. 365–367), nevertheless gives, for transparent plates, the same incorrect formulae as Drude.

In their publications on the polarization of the fluorescence of alkali metal vapors, Wood[12] and Dunoyer[13] use the same incorrect formulae.

Gaviola and Pringsheim[14] point out that because of their neglect of multiple reflections, the values given by Wood are about 13 percent and by Dunoyer about 21 percent too low. To calculate the correct value for four plates they give 28 terms of a doubly infinite series and make the amazing statement (footnote, p. 27):

> Eine allgemeine und nicht sehr komplizierte Formel für die Koefficienten ist wohl nicht anzugeben, wegen der ausserordentlichen Vielfaltigkeit der möglichen Kombinationen. Dagegen lässt sich fur jeden Einzelfall ein relativ einfaches Schema fur die betreffenden Abzählungen leicht aufstellen.

This 103 years after Fresnel had derived the simple formulae of Eq. (1) and they were available in a standard reference book, Winkelmann's *Handbuch der Physik*.

Finally, it may be noted that the correct formulae are given in Schuster, *Theory of Optics*[15] and in Wien-Harms, *Handbuch der Experimental Physik*[16] but the incorrect formulae are perpetuated in O. D. Chwolson *Lehrbuch der Physik*[17] and in Geiger und Scheel *Handbuch der Physik*.[18]

The method which I used in solving the problem of a pile of absorbing plates was simpler than that used by Stokes and gave the results in an entirely different form which, although mathematically equivalent, bears no superficially obvious relationship to his. Stokes' formulae are, so far as I have found, not available in any current text or reference books. For these reasons it seems worth while to present my solution and compare it with his.

The assumptions I made are the same as Stokes made:

> The plates are supposed to be bounded by parallel surfaces, and to be placed parallel to one another. They will also be supposed to be formed of the same material, and to be of equal thickness, except in the case of perfect transparency, in which case the thickness does not come into account. The plates themselves and the interposed plates of air will be supposed, as is usually the case, to be sufficiently thick to prevent the occurrence of the colours of thin plates, so that we shall have to deal with intensities only.

Since the reflection coefficients ρ_{11} and $\rho_{\perp}$ are, in general, different, the results must be calcu-

[8] F. E. Neumann, *Vorlesungen über Theoretische Optik* (Leipzig, 1885), pp. 141–151.

[9] G. G. Stokes, "On the intensity of the light reflected from and transmitted through a pile of plates," Phil. Mag. [4] **24**, 480–489 (1862); Proc. Roy. Soc. London **11**, 545–556 (1862).

[10] Wüllner, *Experimental Physik* (Leipzig, 1899), fifth edition, Vol. 4, p. 727.

[11] Winkelmann, *Handbuch der Physik* (Leipzig, 1906), Vol. 6, p. 1252.

[12] R. W. Wood, "On the emission of polarized light by fluorescent gases," Phil. Mag. **16**, 184–189 (1908).

[13] L. Dunoyer, "Sur la fluorescence des vapeurs des metaux alkalins," le Radium **9**, No. 6, 209–218 (1912).

[14] E. Gaviola and Peter Pringsheim, "Über den Einfluss der Konzentration auf die Polarisation von Farbstoff Lösungen," Zeits. f. Physik. **24**, 24–36 (1924).

[15] Schuster, *Theory of Optics* (New York, 1924), third edition, pp. 51, 52.

[16] Wien-Harms, *Handbuch der Experimental Physik* (Leipzig, 1928), Vol. 18, p. 141.

[17] O. D. Chwolson, *Lehrbuch der Physik* (Braunschweig, 1922), Vol. 2, part 2, pp. 717–718.

[18] Geiger und Scheel, *Handbuch der Physik* (Berlin, 1928), Vol. 19, p. 971.

lated separately for each component. The equations, however, have the same form for both components so that for the rest of the discussion the single symbol ρ will be used.

In passing once through a plate the intensity will be reduced in the ratio of g to 1

$$g = e^{-kT\sec i'}, \tag{3}$$

where k = absorption coefficient of the plates and T = thickness of each plate.

The light reflected by a plate will be made up of that which is reflected at the first surface and that which suffers 1, 3, 5, etc. internal reflections. If the intensity of the incident light be taken as unity the intensities of these various portions will be

$$\rho,\ (1-\rho)^2\rho g^2,\ (1-\rho)^2\rho g^2(\rho^2 g^2),\ (1-\rho)^2\rho g^2(\rho^2 g^2)^2;\ \ldots,\ (1-\rho)^2\rho g^2(\rho^2 g^2)^{(h-1)}, \tag{4}$$

where h is the number of reflections from the rear surface of the plate.

The intensity r of the reflected light from one plate will thus be given by the sum of this geometric series

$$r = \rho + (1-\rho)^2\rho g^2[1-(\rho^2 g^2)^h]/(1-\rho^2 g^2). \tag{5}$$

In all of the references I have found the authors ignore the term $(\rho^2 g^2)^h$ (or for transparent media, ρ^{2h}), tacitly assuming that the number of multiple reflections is infinite. It is possible to show (see Appendix) that in any practical case the term $(\rho^2 g^2)^h$ is so small that no appreciable error is caused by neglecting it. Without that assumption the formulae for more than one plate would be extremely complicated.

Neglecting this term, the formula becomes:

$$r = \rho + (1-\rho)^2\rho g^2/(1-\rho^2 g^2). \tag{6}$$

Similarly the intensity t of the transmitted light will be given by

$$t = (1-\rho)^2 g/(1-\rho^2 g^2). \tag{7}$$

The value of the absorption coefficient and the values of the reflection coefficients for any angle of incidence being known, the corresponding values of r and t can be calculated.

The problem is therefore reduced to the following:

There are n parallel plates of which each reflects and transmits given fractions, r and t, of the light incident upon it: light of intensity unity being incident on the system, it is required to find the intensities of the reflected and refracted light.

Let these intensities be represented by r_n and t_n where r_n and t_n are functions of the plate constants, r and t, and of the variable number, n of the plates. Consider a system of $(m+n)$ plates, and imagine these grouped into two systems of m and n plates, respectively. The incident light, being represented by unity, falling first on the m plates, the light r_m will be reflected from the first group and t_m transmitted. Of the latter fraction, t_n will be transmitted by the second group and r_n reflected. Of the latter fraction t_m will be transmitted by the first group and r_m reflected, etc. Thus for the light reflected by the whole system we have

$$r_{(m+n)} = r_m + t_m^2 r_n + t_m^2 r_m r_n^2 + \cdots t_m^2 r_n (r_m r_n)^{(h-1)}, \tag{8}$$

where h is the number of reflections from the second group.

In this equation also, as in the case of a single plate, it can be proven (see Appendix) that, in any practical case, no appreciable error is caused by assuming an infinite number of reflections. With this assumption Eq. (8) reduces to:

$$r_{(m+n)} = r_m + t_m^2 r_n/(1-r_m r_n). \tag{9}$$

Similarly we find

$$t_{(m+n)} = t_m t_n/(1-r_m r_n). \tag{10}$$

So far my derivation paralleled Stokes' in all essentials. In presenting it I have not hesitated to quote some of his wording almost verbatim, finding no words which would present the thoughts more clearly.

Stokes apparently did not notice that letting $m=1$ in Eqs. (9) and (10) would lead to a simple iteration process for determining r_n. Instead he introduced imaginary angles, α and β, and replaced the integers m and n by continuous variables. By an ingenious process of differentiation, factoring, integrating, and evaluating the integration constants he arrived at an expression of the form:

$$\frac{t_n}{\sin\alpha} = \frac{r_n}{\sin n\beta} = \frac{1}{\sin(\alpha+n\beta)}. \tag{11}$$

Concerning this, Stokes makes the interesting remarks:

> The equations (13), (14) [his numbering] give the following quasi-geometrical construction for solving the problem:—Construct a triangle of which the sides represent in magnitude the intensity of the incident, reflected, and refracted light in the case of a single plate, and then, leaving the first side and the angle opposite to the third unchanged, multiply the angle opposite to the second by the number of plates; the sides of the new triangle will represent the corresponding intensities in the case of the system of plates. I say quasi-geometrical, because the construction cannot actually be effected, inasmuch as the first side of our triangle is greater than the sum of the two others, and the angles are imaginary.

After a page and a half of further transformations he had replaced these complex angles by real variables and gave a final result which may be written:

$$\frac{t_n}{c-c^{-1}}=\frac{r_n}{d^n-d^{-n}}=\frac{1}{cd^n-c^{-1}d^{-n}}, \tag{12}$$

where

$$c=(1-p+q)/r, \tag{13}$$

$$d=(p+q)/t, \tag{14}$$

$$p=(1+t^2-r^2)/2, \tag{15}$$

$$q=(p^2-t^2)^{\frac{1}{2}}=[(1-p)^2-r^2]^{\frac{1}{2}}. \tag{16}$$

Instead of following Stokes, let $m=1$. Then since $r_1=r$ and $t_1=t$, Eq. (6) reduces to

$$r_{(n+1)}=r+t^2r_n/(1-rr_n). \tag{17}$$

Further, writing

$$r_n=ra_n/b_n, \tag{18}$$

where obviously a_n and b_n may be multiplied or divided by any common factor without altering the value of r_n, Eq. (17) becomes:

$$ra_{(n+1)}/b_{(n+1)}=r\left[1+\frac{a_n}{b_n}t^2\Big/\left(1-\frac{a_n}{b_n}r^2\right)\right]. \tag{19}$$

Accordingly we may write:

$$a_{(n+1)}=(1+t^2-r^2)a_n-(a_n-b_n), \tag{20}$$

$$b_{(n+1)}=(1-r^2)a_n-(a_n-b_n), \tag{21}$$

giving

$$a_{(n+1)}-b_{(n+1)}=t^2a_n. \tag{22}$$

By iteration

$$a_{(n+2)}=(1+t^2-r^2)a_{(n+1)}-[a_{(n+1)}-b_{(n+1)}]. \tag{23}$$

Substituting from Eq. (22) gives

$$a_{(n+2)}-(1+t^2-r^2)a_{(n+1)}+t^2a_n=0,$$

or in the notation of Eqs. (15) and (16)

$$a_{(n+2)}-2pa_{(n+1)}+(p^2-q^2)a_n=0. \tag{24}$$

This is a linear recursion formula with constant coefficients. We know, then, that a_n will be given by the coefficient of $x^{(n-1)}$ in the expansion of an expression of the form:

$$\frac{A+Bx}{1-2px+(p^2-q^2)x^2}=a_1+a_2x+a_3x^2+a_4x^3+\cdots. \tag{25}$$

Since $r_1=r$, we may set $a_1=b_1=1$. Substituting these values in Eq. (20) gives:

$$a_2=(1+t^2-r^2)=2p. \tag{26}$$

Substituting these values of a_1 and a_2 in Eq. (25), gives $A=1$, $B=0$. Separating into partial fractions we find

$$\frac{1}{1-2px+(p^2-q^2)x^2}=\frac{p+q}{2q}\cdot\frac{1}{1-(p+q)x}-\frac{p-q}{2q}\cdot\frac{1}{1-(p-q)x}. \tag{27}$$

From this we readily obtain by expansion of both partial fractions:

$$a_n=[(p+q)^n-(p-q)^n]/2q. \tag{28}$$

From Eqs. (22) and (23) we find the two alternative forms:

$$b_n=a_{(n+1)}-(t^2-r^2)a_n=a_n-t^2a_{(n-1)}. \tag{29}$$

Equations (18), (28), and (29) complete the evaluation of r_n.

To evaluate t_n we let $m=1$ in Eq. (10). Substituting for r_n from Eq. (18), this gives:

$$t=\frac{tt_n}{1-rr_n}=\frac{(b_nt_n)t}{b_n-r^2a_n}. \tag{30}$$

From Eq. (29)

$$b_n - r^2 a_n = a_{(n+1)} - (t^2 - r^2)a_n - r^2 a_n \qquad (31)$$
$$= a_{(n+1)} - t^2 a_n = b_{(n+1)}.$$

Thus

$$t_{(n+1)} = (b_n t_n)t/b_{(n+1)}. \qquad (32)$$

Calculating the first two values of b_n and t_n we find

$$b_1 = 1,\ t_1 = t = t^1/b_1,$$
$$b_2 = 1 - r^2,\ t_2 = \frac{t^2}{1-r^2} = \frac{t^2}{b_2}. \qquad (33)$$

By induction we have in general

$$t_n = t^n/b_n. \qquad (34)$$

This completes the solution of the problem.

The form in which these results are expressed bears no superficially evident relation to the form which Stokes derived, but their equivalence is readily shown. Paralleling Stokes' Eq. (12) we may write

$$t_n/t^n = r_n/r a_n = 1/b_n. \qquad (35)$$

Since the equations represent the same mathematical relations we must have

$$\frac{c - c^{-1}}{t^n} = \frac{d^n - d^{-n}}{r a_n} = \frac{cd^n - c^{-1}d^{-n}}{b_n}. \qquad (36)$$

By squaring each of these fractions, calculating $(c-c^{-1})^2$, $(d^n - d^{-n})^2$, and $(cd^n - c^{-1}d^{-n})^2$ from Eqs. (13) and (14), substituting in them $r^2 = (1-p)^2 - q^2$, $t^2 = p^2 - q^2$, and reducing we find

$$\frac{(c-c^{-1})^2}{t^{2n}} = \frac{(d^n - d^{-n})^2}{r^2 a_n^2} = \frac{(cd^n - c^{-1}d^{-n})^2}{b_n^2} = \frac{4q^2}{r^2 t^{2n}}. \qquad (37)$$

Since the numerators and denominators in Eq. (36) are all positive, no ambiguity is introduced by extracting square roots, and the equivalence of the two solutions is confirmed.

For perfectly transparent plates, $t = 1 - r$ and Eqs. (12) and (13) reduce to

$$p = t = 1 - r, \qquad q = 0 \qquad (38)$$

so that Eq. (28) takes on the indeterminate form 0/0. Thus we must write:

$$a_n = \lim_{q \doteq 0}[(p+q)^n - (p-q)^n]/2q$$
$$= \lim_{q \doteq 0}\left(np^{(n-1)} + \binom{n}{3}p^{(n-3)}q^2 + \cdots\right)$$
$$= nt^{(n-1)} \qquad (39)$$

and

$$b_n = a_n - t^2 a^{(n-1)} = nt^{(n-1)} - (n-1)t^n. \qquad (40)$$

So that

$$r_n = r a_n/b_n = nr/[n - (n-1)t]$$
$$= nr/[1 + (n-1)r], \qquad (41)$$

and

$$t_n = (1-r)/[1 + (n-1)r], \qquad (42)$$

which substituting r for ρ, are of the same form as Eqs. (1).

APPENDIX

Upper bound to the error in the calculated intensity of the light reflected from a set of like plates caused by assuming an infinite instead of a finite number of multiple internal reflections between the surfaces.

It is obvious that, other things being equal, the part of the intensity of the reflected light contributed by successive reflections will be greater for fully transparent plates than for plates with appreciable absorption. An upper bound for the error in the case of fully transparent plates will therefore, *a fortiori*, be an upper bound for absorbing plates.

The customary formulae for the fraction r_n of the intensity of the incident light reflected from and transmitted through n surfaces of a set of like parallel transparent plates which take account of multiple internal reflections are equivalent to:

$$r_n = n\rho/[1 + (n-1)\rho], \quad t_n = 1 - r_n = (1-\rho)/[1 + (n-1)\rho], \qquad (1)$$

where ρ, either $\rho_{\parallel}$ or $\rho_{\perp}$, is the reflection coefficient from a single surface.

These are derived by induction from the formula

$$r_{(n+1)}=r_n+(1-r_n)^2\rho[1+\rho r_n+(\rho r_n)^2+\cdots(\rho r_n)^{(h-1)}] \tag{2}$$

$$=r_n+(1-r\rho)^2\rho/(1-\rho r_n)-(1-r_n)^2\rho(\rho r_n)^h/(1-\rho r_n) \tag{3}$$
$$=[\rho+(1-2\rho)r_n]/(1-\rho r_n)-(1-r_n)^2\rho(\rho r_n)^h/(1-\rho r_n),$$

which represents the fraction of intensity coming from h multiple reflections between the set of the first n surfaces and the $(n+1)$th surface.

The customary formulae assume that $(\rho r_n)^h$ is so small that the term containing it in Eq. (4) may be neglected. This is equivalent to assuming that there is an infinite instead of a finite number of these multiple reflections.

The assumption reduces Eq. (4) to

$$r_n=[\rho+(1-2\rho)r_n]/(1-\rho r_n). \tag{5}$$

To estimate the possible relative error involved in this assumption we assume that the errors involved in making the same assumption for any combination of surfaces are so small that their products and squares may be neglected. Since the errors are all positive, the total relative error in $r_{(n+1)}$ will then be the sum of the relative error $\epsilon_{(n+1)}$ in the computation arising from the computation of the first n surfaces and the additional relative error $\zeta_{(n+1)}$ arising from the computation from n to $(n+1)$ surfaces.

In evaluating each partial error we can then neglect the other errors and use the customary formulae for the reflection from any number of plates. If we represent by δr_n the total relative error in the computation of the first n in plates it will produce in $r_{(n+1)}$ a relative error

$$\epsilon_{(n+1)}=\frac{r_n}{r_{(n+1)}}\frac{\partial r_{(n+1)}}{\partial r_n}\delta r_n=\frac{r_n}{r_{(n+1)}}\frac{\partial}{\partial r_n}\frac{\rho+(1-2\rho)r_n}{1-\rho r_n}\delta r_n. \tag{6}$$

Differentiating and substituting values for r_n and $r_{(n+1)}$ from Eqs. (1) gives

$$\epsilon_{(n+1)}=\frac{n}{(n+1)}\cdot\frac{1+(n-1)\rho}{1+n\rho}\cdot\delta r_n<\delta r_n. \tag{7}$$

To evaluate the additional error, $\zeta_{(n+1)}$ in computing from n to $(n+1)$ surfaces we substitute the value of r_n from Eq. (1) in Eq. (4). This gives

$$r_{(n+1)}=\frac{(n+1)\rho}{1+n\rho}-\frac{n^h(1-\rho)\rho^{(2h+1)}}{[1+(n-1)\rho]^h(1+n\rho)}. \tag{8}$$

If we let

$$x=\frac{n^h}{(n+1)}\cdot\frac{(1-\rho)\rho^{2h}}{[1+(n-1)\rho]^h}, \tag{9}$$

we have

$$\zeta_{(n+1)}=x/(1-x). \tag{10}$$

This will give a singular value for

$$d\zeta_{(n+1)}/d\rho=\dot{\zeta}=\dot{x}/(1-x)^2=0. \tag{11}$$

In the physically realizable cases, $0\leqslant\rho\leqslant1$, for both $\rho=0$ and $\rho=1$, $x=0$ and for all values $0<\rho<1$, $0<x<1$. That $0<x$ is obvious. That $x<1$ can readily be shown as follows:

$$n^h(1-\rho)\rho^h<(n+1)n^h\leqslant(n+1)[1/\rho+(n-1)]^h,\quad 0<\rho\leqslant1. \tag{12}$$

Multiplying both sides of the extreme inequality by

$$\frac{1}{(n+1)[(1/\rho)+(n-1)]^h}=\frac{\rho^h}{(n+1)[1+(n-1)\rho]^h} \tag{13}$$

gives $x<1$. Then $\dot{x}=0$ must correspond to a maximum of both x and $\zeta_{(n+1)}$ with $0<\rho<1$. Differentiating and rearranging give for the maximum:

$$-(n+1)[1+(n-1)\rho]^{(h+1)}\dot{x}/n^h\rho^{(2h-1)}=(n-1)(h+1)\rho^2-[(n-3)h-1]\rho-2h=0. \tag{14}$$

For $n=1$ this reduces to a linear equation giving:

$$\rho=2h/(2h+1)<1. \tag{15}$$

For $n>1$ the positive root is given by:

$$\begin{aligned}\rho&=\{+[(n-3\cdot h-1)^2+8(n-1)h(h+1)]^{\frac{1}{2}}+(n-3)h-1\}/2(n-1)(h+1)\\&=\{+[\overline{(n+1}\cdot h-1)^2+8nh]^{\frac{1}{2}}+(n-3)h-1\}/2(n-1)(h+1)\\&=\{+[\overline{(n+1}\cdot h+3)^2-8(h+1)]^{\frac{1}{2}}+(n-3)h-1\}/2(n-1)(h+1).\end{aligned} \tag{16}$$

From these we have:

$$0<\frac{(n-1)h-1}{(n-1)(h+1)}=\rho_0<\rho<\rho_1=\frac{(n-1)h+1}{(n-1)(h+1)}\lesseqgtr 1 \tag{17}$$

We wish to know how, for constant h, the maximum value of $\zeta_{(n+1)}$ varies with n. Differentiating x with respect to n subject to the conditions $h=$constant and

$$\dot{x}=\partial x/\partial\rho\,|_{n,h}=0$$

we have:

$$x'=\left.\frac{\partial x}{\partial n}\right|_h=\left.\frac{\partial x}{\partial n}\right|_{h,\rho}+\left.\frac{\partial x}{\partial\rho}\right|_{n,h}\left.\frac{\partial\rho}{\partial n}\right|_h=\left.\frac{\partial x}{\partial n}\right|_{h,\rho} \tag{18}$$

$$=\frac{(1-\rho)\rho^{2h}n^{(h-1)}}{(n+1)^2[1+(n-1)\rho]^{(h+1)}}\{(n+1)h-n-[n(n-1)+(n+1)h]\rho\}. \tag{19}$$

The factors outside of the brackets are all positive so that the sign of x' and consequently of $\zeta'_{(n+1)}$ will depend upon the sign of the quantity enclosed in the brackets. For $n=1$ we have, for maximum error, $\rho=2h/(2h+1)$ and the quantity in the brackets reduces to

$$2h-1-2h\cdot\frac{2h}{2h+1}=[(2h-1)(2h+1)-4h^2]/(2h+1)=-1/(2h+1)<0. \tag{20}$$

For constant n and h the quantity in the brackets is a monotone decreasing function of ρ. If, for any value of ρ it is negative, it will be negative for all larger values of ρ. Substituting in it the lower boundary of ρ, ρ_0 from Eq. (17) gives:

$$(n+1)h-n-[n(n-1)+(n+1)h]\frac{(n-1)h-1}{(n-1)(h+1)}=\frac{1+n-n^2}{h+1}. \tag{21}$$

For $n\geqslant 2$, $h>0$ this is always negative. Thus for all values of $n\geqslant 1$, $h>0$ we have for values of ρ corresponding to a maximum error

$$x'=\frac{\partial x}{\partial n_h}<0. \tag{22}$$

We conclude then, that the maximum relative error involved in computing, by Eqs. (5) for a given number of multiple reflections, from n to $(n+1)$ surfaces, is less than the relative error involved in computing from one to two surfaces for the same number of multiple internal reflections.

For $n=1$ Eq. (9) reduces to

$$x=\tfrac{1}{2}(1-\rho)\rho^{2h}, \tag{23}$$

giving

$$\delta r_2=\zeta_2=\frac{(1-\rho)\rho^{2h}}{2-(1-\rho)\rho^{2h}}. \tag{24}$$

Substituting the value of ρ for maximum error from Eq. (15) gives

$$\delta r_2=\zeta_2=\frac{[2h/(2h+1)]^{(2h+1)}}{4h-[2h/(2h+1)]^{(2h+1)}},$$

which is then an upper bound to the error involved in calculating r_2 by Eq. (5) instead of Eq. (4)

It is convenient to write this:

$$1/4h\delta r_2=(1+\tfrac{1}{2}h)^{(2h+1)}-1/4h. \tag{25}$$

Differentiating this with respect to h we find

$$\frac{d}{dh}(1/4h\delta r_2)=-(2h+1)/2h^2+1/4h^2=-(4h+1)/4h^2. \tag{26}$$

Thus $1/4h\delta r_2$ is a monotone decreasing function of h for all $0<h<\infty$. For $h\doteq\infty$ we have:

$$\lim_{h\doteq\infty}(1/4h\delta r_2)=e=2.71828+. \tag{27}$$

Then for all $0<h<\infty$

$$1/4h\delta r_2>e, \tag{28}$$

from which $\delta r_2<1/4eh<0.09197/h$. This is therefore *a fortiori* an upper bound to the error in question. It is furthermore, in view of inequality (22), *a fortiori*, an upper bound to the error involved in computing from n to $(n+1)$ surfaces for the same value of h.

Taking into consideration inequality (7) we can finally state:

$$\delta r_n<(n-1)/4eh<0.09197(n-1)/h,$$

where h is the minimum number of internal reflections for any combination of the surfaces.

It would only approach that value near grazing incidence. Thus for 10 reflections from the back surface of a single plate the maximum possible error is less than 0.92 percent and is given by $\rho=0.9254$. With $\mu=1.50$, for ρ_{11} this gives by Fresnel's formulae, $i=88°45'$ and for $\rho_{\perp}$, $i=89°27'$. For the polarizing angle 56°19′, at which the plates are ordinarily used

$$\rho_{11}=(\mu^2-1)^2/(\mu^2+1)^2=0.14793.$$

Here the second reflection cannot contribute as much as 1 percent and the third as much as 2/100 percent of the total reflected light.

It is obvious that in any practical case the assumption of an infinite number of reflections does not introduce any appreciable error.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Optical Scattering by Dielectric Powders of Uniform Particle Size*

J. R. DEVORE
Research Division, Technical Department, The New Jersey Zinc Company, Palmerton, Pennsylvania

AND

A. H. PFUND
Department of Physics, Johns Hopkins University, Baltimore 18, Maryland

(Received May 28, 1947)

The spectral-transmission curve of a film of dielectric powder having uniform particle size shows a very pronounced minimum corresponding to the Mie scattering maximum. The wave-length of this "Mie minimum" changes as the refractive index of the medium surrounding the particles is changed. If a film of dielectric powder is prepared and the wave-lengths of the Mie minima are determined when liquids of different refractive index are flowed around the particles then, if the wave-length values are plotted against the refractive indices of the surrounding media, the points fall on a straight line that intersects the axis at the refractive index of the powder, thereby providing a method for measuring the refractive index of powders above the range of immersion liquids. A method, using the data for the Mie minima, is described for determining the optimum particle size for white paint pigments. This method is illustrated with data on zinc sulfide and titanium dioxide.

AN unpublished dissertation by R. L. Henry contains spectral-transmission data of some films of uniform particle-size zinc sulfide. These transmission curves show a very pronounced minimum corresponding to the Mie scattering maximum. One of the authors (A.H.P.) observed the behavior of the "Mie minimum" as the medium surrounding the zinc sulfide particles is changed to one of different refractive index, and started a program to study the phenomenon further. The results of the study led to empirical relationships between this behavior of the "Mie minimum," the refractive index of the particles, and the particle size for maximum opacity. Subsequently, these relationships were confirmed theoretically.

EXPERIMENTAL

For this work it is necessary to have a powder of very uniform particle size, which can be obtained by elutriation if the original sample is non-uniform. A powder film is prepared from the sample by the following procedure. A suitable amount (ranging from 0.002 gram to 0.008 gram) of sample is weighed out, placed in a glass mortar and dispersed in benzene with the aid of a wetting agent such as poppyseed oil if necessary. In the meantime, a microscope cover glass (1" square) has been cleaned and placed in the center of the bottom of a 200-cc centrifuge beaker and covered with about 20 cc of clear benzene. The slurry in the mortar is emptied carefully into the beaker without disturbing the cover glass on the bottom. The mortar is washed with additional benzene until the beaker contains from 150 to 175 cc. Then the slurry in the beaker is gently stirred; the beaker is covered to prevent evaporation, placed in a spot where it will not be exposed to excessive temperature changes, and allowed to remain undisturbed for at least four days. At the end of this time the pigment will have settled out on the cover glass and bottom of the beaker and the brownian motion of the liquid will have packed the settled pigment firmly on to the plate. After the clear benzene is drawn off so carefully that the pigment deposit on the cover glass is not disturbed, the small amount of benzene remaining in the beaker is evaporated. The plate with the dry powder film is made into a cell by placing a tinfoil ring on the surface of the film and clamping a clear cover glass on top. This cell is mounted in a spectrophotometer along with a clear cell. If the pigment sample should be so coarse that the Mie minimum occurs in the infra-red at a wave-length longer

* This article was submitted by the senior author without the knowledge of Professor Pfund that it was intended for a special issue of the Journal.

than 4 microns, it will be necessary to use rocksalt plates instead of microscope cover glasses.

The spectral transmission of the powder film is first measured with air in the cells. Then the cells are filled (by capillarity) with a liquid of known refractive index and the spectral transmission is measured again. This liquid is evaporated from the cells by placing in an oven if necessary and another liquid is introduced into the cells and the spectral transmission is measured. This process is repeated as often as desired.

The first sample measured in this procedure was a sample of luminescent-grade zinc sulfide having uniformly large particle size. The sample was measured microscopically and the values obtained were $d_1 = 0.86$ micron, $d_3 = 1.07$ micron. (d_1 is the arithmetic mean diameter; d_3 is the specific surface mean diameter. The difference between these two values is a measure of the non-uniformity of the sample.) The dry powder film of this sample contained 0.001-gram pigment on the 1 square inch of surface. The transmission curves, shown in Fig. 1A, were measured with a Gaertner rocksalt prism, infra-red spectrophotometer when the medium surrounding the zinc-sulfide particles was air ($n_0 = 1.00$), methyl alcohol ($n_0 = 1.33$), benzene ($n_0 = 1.49$), and chlorophanthalene ($n_0 = 1.63$). It is evident that this sample is sufficiently uniform in particle size to make the Mie minimum quite pronounced and easily located. Similar but less pronounced minima were shown in transmission curves previously published by one of the authors.[1]

If the wave-length at which each of these

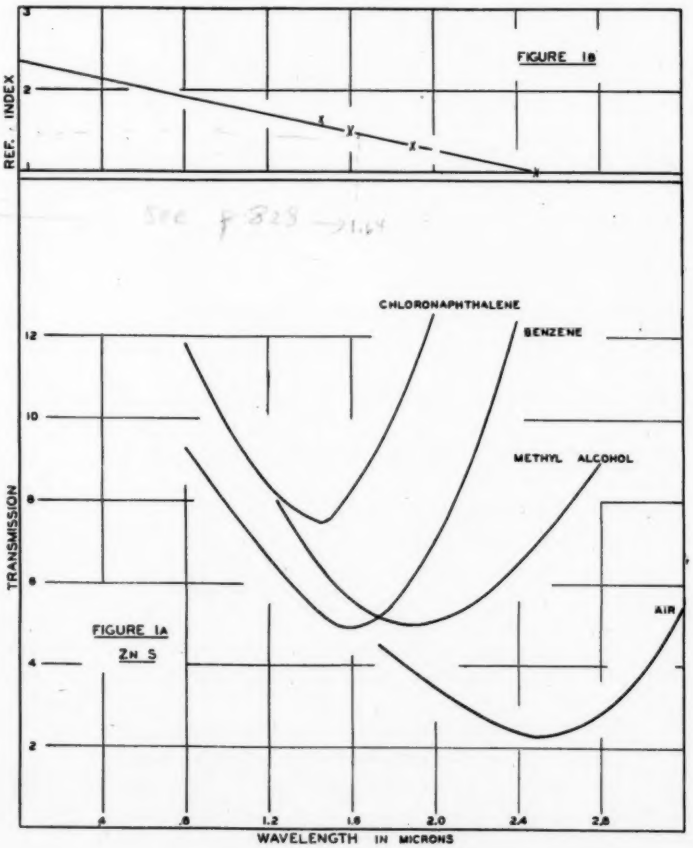

FIG. 1.

[1] A. H. Pfund, J. Opt. Soc. Am. 29, 10 (1939).

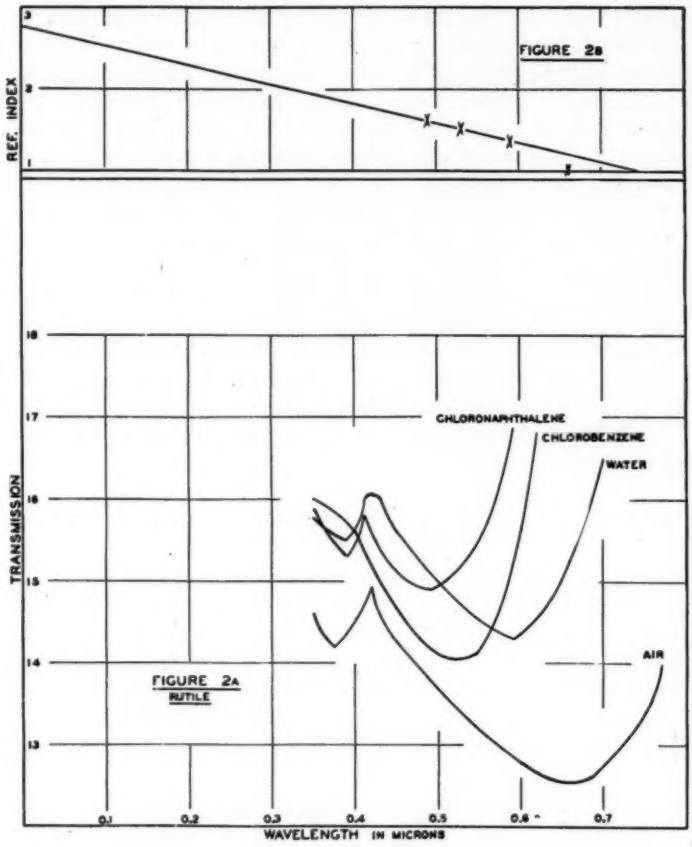

FIG. 2.

minima occurs is plotted against the refractive index of the surrounding medium the results are shown in Fig. 1B. The points fall on a straight line and if this line is extrapolated to $\lambda = 0.0$ it intersects the axis at approximately $n_0 = 2.37$, which is the refractive index of zinc sulfide. This line enables one to find the wave-length at which the Mie minimum will occur when the refractive index of the surrounding medium is known. In particular, when the surrounding medium is linseed oil ($n_0 = 1.48$) the Mie minimum occurs at $\lambda = 1.64$ micron. If a sample of ZnS were of such a size that it would have maximum opacity when made into a paint in linseed oil it would have the Mie minimum located at about 0.5 micron. It will be shown in the theoretical section that the diameters of the desired sample and the experimental sample are in the same ratio as the wave-lengths of the Mie minima, so one can calculate the diameter of a sample of ZnS having maximum opacity

$$d_3 \text{ (optimum)} = 0.5/1.64(1.07) = 0.33 \text{ micron.}$$

A sample of pigment-grade rutile TiO_2 which measured $d_1 = 0.39$ micron, $d_3 = 0.41$ micron was examined in the same manner. The dry powder film of this TiO_2 contained about 0.0006 gram of pigment on the 1 square inch of surface. The transmission curves shown in Fig. 2A were measured with a Beckman spectrophotometer when the surrounding medium was air ($n_0 = 1.00$), water ($n_0 = 1.33$), chlorobenzene ($n_0 = 1.52$), and chloronaphthalene ($n_0 = 1.63$). These transmission curves, with the exception of the one for chlorobenzene, show secondary minima at shorter wave-lengths that resemble the secondary peak

in the Mie-scattering curve. However, as the refractive index of the surrounding medium increases, the secondary minimum shifts slightly toward longer wave-lengths rather than toward shorter wave-lengths as would be expected if they were simply secondary Mie minima. Consequently, the cause of these secondary minima is not known.

Figure 2B shows the wave-length of the Mie minimum plotted against the refractive index of the surrounding medium. A straight line intersecting the axis at $n_0 = 2.75$, which is the approximate value of the index of rutile, passes through three of the four points. It is possible that the deviation of the fourth point is a result of the large optical dispersion of rutile in the visible spectrum.

Following the same procedure used with the sample of ZnS the optimum diameter of rutile particles can be calculated. In this sample, when $n_0 = 1.48$, the Mie minimum occurs at 0.54 micron, so

$$d_3 \text{ (optimum)} = 0.5/0.54(0.41) = 0.38 \text{ micron.}$$

A sample of classified silica was examined by the same procedure. The dry powder film of this material contained about 0.0022 gram on 1 square inch. The transmission curves shown in Fig. 3A were measured in air ($n_0 = 1.00$), water ($n_0 = 1.33$), and Cellosolve ($n_0 = 1.41$). An attempt was made to measure the transmission in chloronaphthalene but any phenomenon such as the Mie minimum occurred at wave-lengths too short to be measured.

Figure 3B contains the graph of the wave-length of the Mie minimum plotted against n_0 as before. When the straight line through the three points is extrapolated, it intersects the axis at

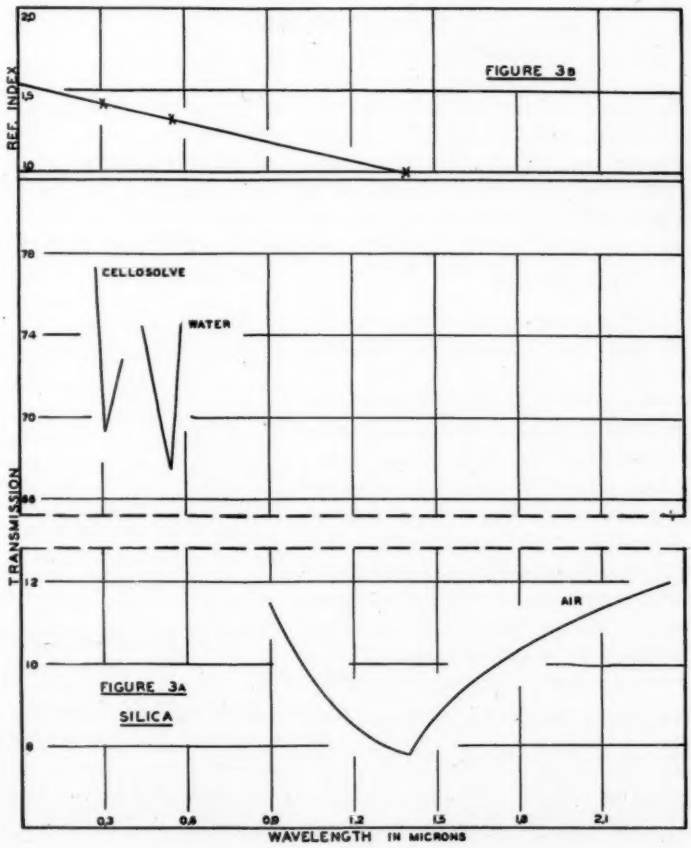

FIG. 3.

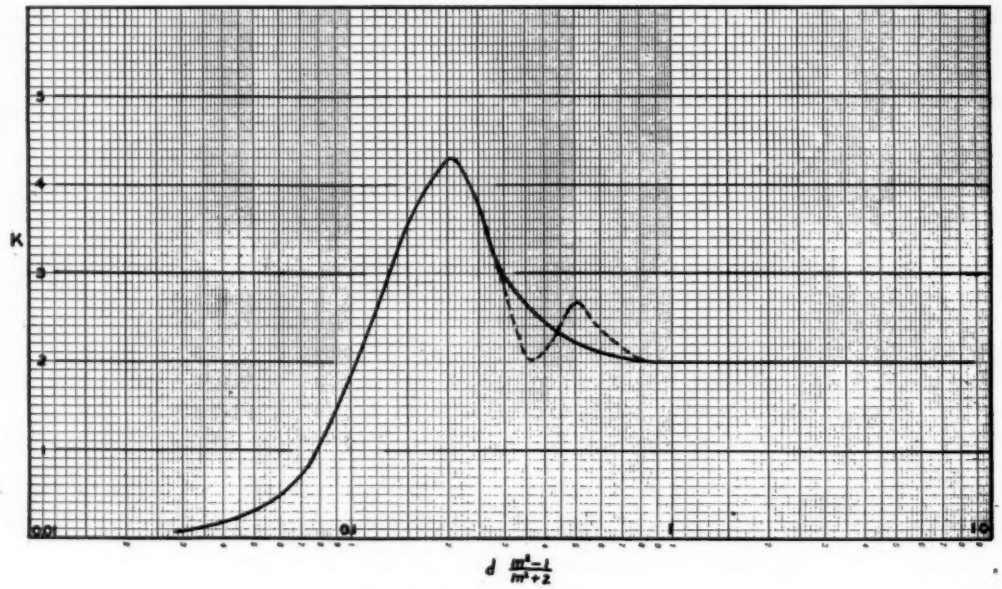

FIG. 4. Universal scattering curve.

$n=1.53$ which was proved by immersion method to be the index of the silica sample. The particle size for maximum opacity was not calculated because the refractive index of this material is so low that it is of no use as an opacifying pigment.

THEORETICAL

Stratton and Houghton,[2] and more recently Barnes and LaMer[3] have published theoretical curves based on the equations of Mie[4] and Debye,[5] for the optical scattering of dielectric particles whose size is the same order of magnitude as the wave-length of light. These same authors published experimental curves showing that the theoretical curves are verified experimentally as far as it is possible experimentally to duplicate the theoretical conditions. These theoretical scattering curves are specific for the relative refractive indices of particle and surrounding medium for which they were calculated. However, Bailey[6] found empirically that if the ratio of particle size to wave-length (d/λ) is multiplied by the Lorentz refractive-index coefficient, m^2-1/m^2+2, where m is the relative index of refraction, then the resultant scattering curve is independent of the refractive index of the particles or the medium. Such a universal scattering curve has been published in a report by LaMer and others[7] and is reproduced in Fig. 4.

The transmission curve for a film of dielectric powder having a perfectly uniform particle size, will be the mirror image of the scattering curve shown in Fig. 4. Consequently, the transmission curve will have a rather sharp and deep minimum which has been designated in this article as the Mie minimum.

Now it is necessary to study the behavior of the Mie minimum as the conditions of measurement are changed.

1. Change of Refractive Index of the Surrounding Medium

It is evident from the universal scattering curve that $(d/\lambda)(m^2-1/m^2+2)$ will have a con-

[2] J. A. Stratton and H. G. Houghton, Phys. Rev. **38**, 195 (1931).
[3] M. D. Barnes and V. K. LaMer, J. Coll. Sci. **1**, 79 (1946).
[4] G. Mie, Ann. d. Physik **25**, 377 (1908).
[5] P. Debye, Ann. d. Physik **30**, 57 (1909).
[6] E. D. Bailey, Ind. Eng. Chem. (Anal. Ed.) **18**, 365 (1946).
[7] V. K. LaMer, J. Q. Umberger, D. Sinclair, and F. E. Buchwalter, U. S. Office of Technical Services, Report PB32208.

stant value at the Mie minimum, so

$$(d/\lambda)(m^2-1/m^2+2)=k,$$

or

$$(d/\lambda)(n^2-n_0^2/n^2+2n_0^2)=k, \qquad (1)$$

where n = refractive index of dielectric powder, n_0 = refractive index of surrounding medium. Under the conditions being considered, d and n are constant. For convenience, Eq. (1) may be written as

$$(k/d)\lambda=(n^2-n_0^2/n^2+2n_0^2). \qquad (1')$$

In this equation λ is the wave-length at which the Mie minimum occurs. Starting with a low value of n_0 and increasing, it is evident that the value of λ decreases until at $n_0=n$, $\lambda=0$. That is, the minimum shifts toward shorter and shorter wave-lengths until when $n_0=n$ it disappears. Now in order to study the way λ varies with n_0, the right side of Eq. (1') can be expanded in a series

$$(k/d)\lambda=1-3(n_0/n)^2+6(n_0/n)^4-12(n_0/n)^6. \qquad (2)$$

If one plots the curves for the functions

$(k/d)\lambda=1-3(n_0/n)^2$,
$(k/d)\lambda=1-3(n_0/n)^2+6(n_0/n)^4$,
$(k/d)\lambda=1-3(n_0/n)^2+6(n_0/n)^4-12(n_0/n)^6$,

etc., formed by adding successive terms of the series in Eq. (2), the following results become apparent:

(a) For all values of n_0 between 0 and n the curves fall alternately on one side and the other of a straight line from $(k/d)\lambda=1$ at $n_0=0$ to $(k/d)\lambda=0$ at $n_0=n$.

(b) For small values of n_0 the values of $(k/d)\lambda$ approach this straight line as the number of terms increases.

(c) For large values of n_0 the absolute values of $(k/d)\lambda$ become progressively larger as the number of terms increases but they oscillate about this straight line as a median value.

Thus this straight line becomes the limit for this series over the range $0 \leqq n_0 \leqq n$, so the behavior of the experimental points is expected. It has been assumed in this work that the refractive index, n, of the dielectric is constant. This is nearly true for many dielectrics over the spectral range in which the measurements are made. However, some materials such as rutile exhibit strong dispersion, and it would then be expected that the points would not fall on a straight line. This failure to fall on a line is shown by the data for the rutile sample.

2. Change of Particle Size of the Dielectric

Now in Eq. (1) the quantities n and n_0 are constant and d and λ are the variables. One obtains the simple linear equation

$$n^2-n_0^2/n^2+2n_0^2d=k\lambda, \qquad (1'')$$

showing that λ varies directly with d. This relationship can be used to determine the particle size (denoted by d_m) that will produce maximum opacity with a given white pigment material. This procedure was illustrated in the experimental section. It follows from Eq. (1'') that when n_0 is constant

$$d_1/d_2=\lambda_1/\lambda_2,$$

where λ_1 is the wave-length of the Mie minimum for a pigment sample of size d_1 and λ_2 is the wave-length of the minimum for a sample of the same material but having a size d_2. In particular

$$\frac{d_m}{d_{\text{sample}}}=\frac{\lambda_m}{\lambda_{\text{sample}}}=\frac{0.5}{\lambda_{\text{sample}}}, \qquad (3)$$

where λ_{sample} is the value of λ for the given sample when $n_0=1.48$. The value 0.5 micron was substituted for λ_m since a sample of maximum opacity will have the Mie minimum in the visible green. Equation (3) is the equation that was used in the experimental section to calculate the particle size having maximum opacity.

SUMMARY AND CONCLUSIONS

Some interesting observations have been presented on the spectral transmission of films of dielectric powders having uniform particle size. In particular, a study was made of the behavior of the Mie minimum when the refractive index of the medium surrounding the particles is changed. It was found that when the wave-length of the Mie minimum is plotted against the refractive index of the surrounding medium, the points fall on a straight line that intersects the axis at the refractive index of the powder. This provides a method for measuring the refractive index of a powder that is outside the range of immersion liquids, but the method is subject to the following limitations.

(a) The sample must have very uniform particle size, which means, in most cases, that the sample must be classified by elutriation or some other means.

(b) The particle size must be such that the Mie minimum will fall in a measurable region of the spectrum.

(c) For a high index material, a long extrapolation is involved and consequently is subject to a moderate amount of error.

The study has provided a method for obtaining a value of the diameter of the particles having maximum opacity for a given pigment material. Not much can be said about these measurements at present. No direct experimental determination of the particle size for maximum opacity has been carried out for the various pigment materials so it is possible to say only that the results are about the right order of magnitude. Because of the difficulty of obtaining samples having sufficient uniformity of particle size we have not made the determination of optimum particle size from more than one sample of the same material so we have no indication of the uniqueness of the value determined in this manner. When we compare the values obtained for ZnS and rutile TiO_2, it appears that the value determined this way cannot be very accurate because the value for rutile is larger than for ZnS. According to the theory presented above, the optimum diameter decreases as the refractive index of the pigment increases.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Optical Study of Acoustic Fields Near Diffracting Edges*

J. C. HUBBARD, I. F. ZARTMAN, AND C. R. LARKIN
Radiation Laboratory, The Johns Hopkins University, Baltimore, Maryland

(Received July 18, 1947)

Diffraction phenomena in acoustic fields of ultrasonic frequency in air have been photographed by schlieren technique and by shadow methods employing a spark of extremely short duration. Variations of intensity in the field are observed in the immediate vicinity of apertures and obstacles as indicated by theoretical treatment of the problem in a more rigorous manner than that of Kirchhoff.

THE theory of diffraction phenomena involves a solution of the wave equation $\nabla^2\phi+k^2\phi=0$, which is subject to certain boundary conditions and, in addition, is restricted by other relations such as Maxwell's equations for electromagnetic waves or an equation of continuity for pressure waves. The classical treatment of the problem has been the evaluation of Kirchhoff's integral expression which is based upon the assumption that the field in the aperture is not influenced by the presence of the diffracting screen. That this assumption leads to inconsistent results was pointed out as long ago as 1896 by Sommerfeld[1] who worked out a more rigorous treatment for the case of the semi-infinite plane with straight diffracting edge. More recently, Morse and Rubenstein[2] have developed expressions applicable to the case of a slit (or of a ribbon) and have called attention to the existence of suitable tables for their numerical reduction.

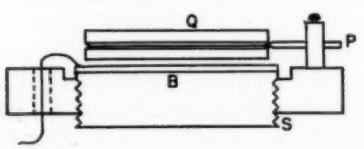

FIG. 1. Schematic detail of crystal mounting.

While the Kirchhoff method yields results in fair agreement with observation under the so-called Fraunhofer and Fresnel conditions, it is totally inadequate for treating the region in the immediate vicinity of the slit itself.

* This work was supported in part by Contract NOrd 8036 with the Navy Department.

[1] Sommerfeld, Math. Ann. **47**, 317 (1896).
[2] Morse and Rubenstein, Phys. Rev. **54**, 895 (1938).

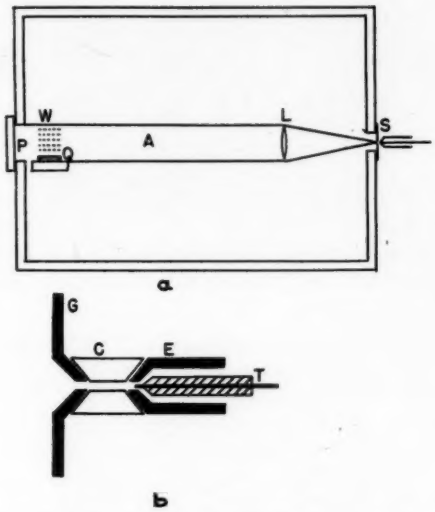

FIG. 2. a. Box for producing spark-shadow pictures. b. Detail of spark gap.

The wave-lengths of microwaves and of ultrasonic waves are of such values that one may produce diffraction patterns in and close to apertures of convenient size for experimental study. The recent publication of Andrews[3] on the measurement of the diffraction pattern of microwaves in a circular aperture has appeared during the progress of this work.

The present study was undertaken with the object of determining the pattern of amplitude of vibration in the immediate vicinity of a diffracting edge or slit, using for the purpose ultrasonic waves in air of 0.7-mm wave-length, and with slits from one-quarter to five wave-lengths in width. The results may be compared with those which have been obtained for the radiation into liquids from piezoelectric crystals of small radiating area.[4]

The first requirement of such a study is a system of truly plane waves of constant frequency and of unvarying intensity at the source. Such a system is provided by a circular x-cut crystal of untwinned quartz supported by pins at three equidistant points on the circumference. In making provision for this mounting, a very shallow V-cut is made around the edge of the crystal with its vertex accurately in the median plane. Pins with conical points, the angle of which is somewhat less than the angle of the

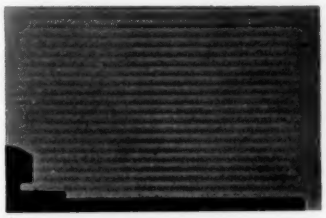

FIG. 3. Instantaneous spark-shadow picture of crystal radiating into free space.

V-cut are held by supports so as to press gently into the V-groove. The quartz plate is thus supported in an unvarying position. A backing plate[5] B is adjusted at a quarter-wave distance behind the quartz plate, effectively increasing the radiation from the other surface. The mounting arrangement is shown in Fig. 1.

By use of a spark of short duration such as developed by Beams and Snoddy, *et al.*,[6] it is possible to take an almost instantaneous shadow photograph of the sound field so that dark bands on the negative correspond to condensations in the sound wave, the dark bands thus being separated by one *whole* wave-length. This is accomplished in the following manner. A felt-lined box, $24\times16\times16$ inches, having a pinhole at the center of one end, allows the flash of light of approximately 10^{-7} sec. duration to be made into a parallel beam by the objective L, Fig. 2a. This parallel beam crosses the sound-wave system W, the condensations and rarifactions of which cause a retardation of the phase of the light in the more dense layers and similarly an advancement in the less dense layers of the sound field. There thus results on the negative, at P, a distribution of density in which dark bands are one whole wave apart. To test this method a crystal of 494.0-Kc frequency of thickness vibrations was mounted so that the radiating surface of the quartz was accurately parallel to the optical axis of the photographic system *SLP*. When the crystal was driven at its natural fre-

[3] Andrews, Phys. Rev. **71**, 777 (1947).
[4] Compare, e.g., Osterhammel, Akustische Zeits. **6**, 73 (1941).
[5] Alleman, Dissertation, Johns Hopkins University, 1938; Phys. Rev. **55**, 87 (1939).
[6] Beams, Kuhlthau, Lapsley, McQueen, Snoddy, and Whitehead, J. Opt. Soc. Am. **37**, 868 (1947).

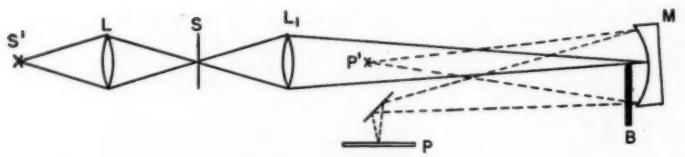

FIG. 4. Diagrammatic arrangement for obtaining schlieren photographs.

quency of thickness vibration the photograph, shown in part in Fig. 3, was obtained; the temperature of the room being close to 27°C. A study of the negative shows that the wave system is composed of plane, equally spaced waves. Measurement of the first 41 band centers gives a value 0.7021 ± 0.0030 mm as the wave-length at 27°C. This gives a velocity of 346.8 m/sec. or, reduced to 0°C, 330.8 m/sec., in excellent agreement with values obtained from open air by methods of greater precision.

The earliest schlieren photograph of standing ultrasonic waves in air known to the authors, was made by Pohlmann[7] who used a modification of the Toepler method. By a similar method, Alleman[5] was able to photograph progressive sound fields, and standing waves in gases under various conditions.

The arrangement used in the present study is shown in Fig. 4. Light from a mercury arc with a green filter is concentrated on a slit S and focused by lens L_1 on the edge of a stop B, which covers half the 3-meter, 6-inch aperture mirror, M. The sound-wave system is placed in the light beam near L_1, and is adjusted in position so as to be at a distance from the mirror equal to its radius of curvature. The real image, of natural size, of the wave system is formed in the plane P, and may be viewed by an eyepiece or formed directly on film in a camera box. Viewed in steady illumination, the dark field is lighted only by such light as is deviated by the sound field. One thus obtains a density of negative, depending

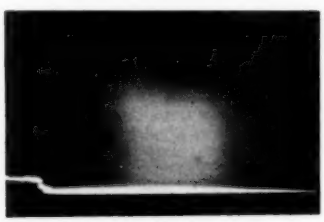

FIG. 5. Schlieren photograph of sound field from crystal radiating into free space.

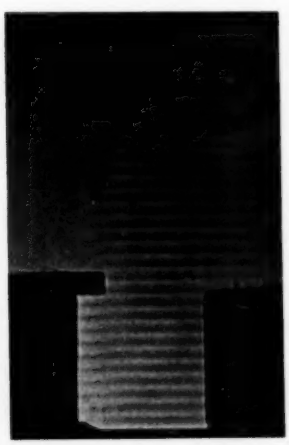

FIG. 6. Diffraction of sound field at straight edge (spark-shadow graph).

FIG. 7. Diffraction at straight edge (schlieren).

FIG. 8. Diffraction by slit 5λ in width (spark-shadow graph).

[7] Pohlmann, Naturwiss. **23**, 511 (1935).

FIG. 9. Diffraction by slit 2λ in width (spark-shadow graph).

FIG. 10. Diffraction by slit λ in width (spark-shadow graph).

FIG. 12. Diffraction by slit 5λ in width (schlieren).

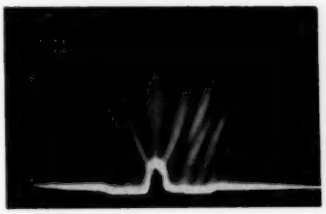

FIG. 13. Diffraction by cylinder (wire) $\frac{3}{4}\lambda$ in diameter and 2λ above crystal (schlieren).

upon density variations of the medium in the sound field, i.e., upon amplitude gradiant in the sound wave. In order to photograph standing waves it is necessary to stop the light of the central image only, so that diffracted light from both sides of the mirror may fall upon the plate.

When the system producing the progressive waves shown in Fig. 3 is transferred to the schlieren camera, the sound field shown in Fig. 5 is seen. As would be expected, the intensity of light is greatest when looking over the center of the circular crystal, falling away as the light path through the sound field decreases to zero over the edges of the circular area.

When a straight edge is formed intercepting the sound field over the middle of the crystal, the spark-shadow picture, Fig. 6, is obtained. In this case the straight edge was a sheet 5/4 λ thick, at about $10\frac{1}{4}$ λ for the face of the crystal. A comparison may be made with the schlieren picture, Fig. 7, in which, however, the distance to crystal face was λ/4. Light scattered at the surfaces of the crystal and straight edge tends to obscure detail, but structure in the system of streamers may be observed into the region between crystal and diffracting plane.

Figures 8 to 11 are spark-shadow photographs of slits 5, 2, 1, and $4\frac{1}{2}\lambda$ in width, respectively, the slit being λ/4 above the crystal in each case. Figure 11 is marred by a bad scratch and also

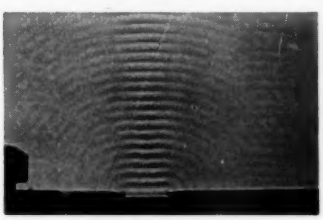

FIG. 11. Diffraction by slit $4\frac{1}{2}\lambda$ in width (spark-shadow graph).

FIG. 14. Diffraction by cylinder 4/3λ in diameter and 2λ above crystal (schlieren).

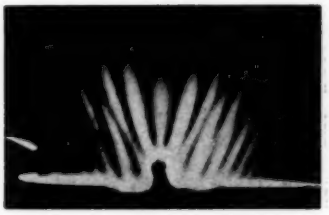

FIG. 15. Diffraction by cylinder 4/3λ in diameter and 2¼λ above crystal (schlieren).

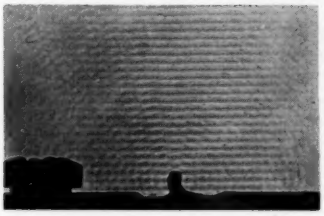

FIG. 16. Diffraction by cylinder 4/3λ in diameter and 2¼λ above crystal (spark-shadow graph).

by leakage of plane waves through a hole in the razor blade which was used for slit material, but is included because of the good contrast in the diffraction system on the left of the aperture. Figure 12 is the schlieren counterpart of Fig. 8; the structure being observed almost to the crystal surface, and the non-uniformity of the field in the plane of the aperture clearly observable.

Figures 13 to 15 are schlieren pictures produced by a wire supported in a plane parallel to the crystal face and so as to be seen axially in the photographs. The wire diameter used for Fig. 13 was ⅔ λ, while that for Figs. 14 and 15 was 4/3 λ. In Figs. 13 and 14 the near surface of the wire was 2 λ above the crystal, while in Fig. 15 this distance was 2¼ λ. Figure 16 is the spark picture with conditions identical to those of Fig. 15, and shows the wave system corresponding to the streamers of the schlieren picture.

The photographs exhibited are selections from some obtained in a survey to explore the possibilities of the methods described above for use in a more extended study of acoustic diffraction and related problems.

The authors wish to acknowledge the assistance of Mr. Earl Byerly and Mr. Herbert Seidel of this Laboratory who have constructed the apparatus and taken the photographs.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Fluctuations in Gas Discharges

H. Y. Loh and G. H. Dieke
The Johns Hopkins University, Baltimore, Maryland

(Received May 29, 1947)

Examination of the light output of discharge tubes, using the multiplier-type photo-tube in conjunction with a cathode-ray oscillograph, is suggested as a powerful aid to the usual electrical measurements in understanding the behavior of electrical discharges in gases. Exhaustive experiments of this type are under way, the preliminary results of which are reported in this paper. A striking characteristic of these results is the appearance of oscillations not only in the voltage and current, but particularly in the light intensity. These oscillations appear to be very general, and are not due to the external circuit used. They are only of a few percent magnitude in the case of the voltage and current, but often nearly 100 percent in the case of the light intensity. It seems certain that the theory of electrical discharges should in many cases be revised to take these oscillations into account.

1. INTRODUCTION

ALTHOUGH much work has been done in the past on phenomena occurring in gas discharges,[1] there are quite a number of unsolved problems concerning the mechanism of such discharges. For this reason, any additional information is welcome which can supplement the data obtained from the well-established methods of investigating gas discharges. These are chiefly current and potential measurements, often done with the help of probes.

Practically all gas discharges give off light. The distribution of this light in the various parts of a discharge is quite characteristic, and often has been used to illustrate the discharge mechanism. On the other hand, it appears that when properly examined these light phenomena are not quite as simple as is usually assumed. After this had been established in a few casual observations in our laboratory, a more systematic survey was made of the light phenomena occurring in a variety of types of gas discharges. The results of this survey are presented in the present paper. It will be seen that the phenomena are rather complex, and that a complete understanding in terms of the discharge mechanism will require additional work under more carefully controlled conditions. Such investigations are under way in our laboratory and have already yielded significant results which will be reported in a subsequent paper.

Most quantitative investigations on discharge phenomena have been carried out in the past under stationary conditions. The special phenomena of spark-like character occurring when the discharge is first started are then passed over. A complete study of steady discharges is rather tedious, as voltage and current must be varied systematically, and measurements made at each set of conditions. In many respects the study of an alternating-current discharge easily yields much more detailed information even though the data are usually less precise. For moderate frequencies the discharge can be regarded as a succession of steady discharges with varying voltages applied to the discharge tube and, therefore, observations within one half-cycle will yield data that can only be laboriously obtained with a set of systematically changed steady discharges. In addition the discharge must be reignited every half-cycle and, therefore, the transient ignition phenomena can also be studied at ease. For a general survey the study of a.c. discharges has therefore definite advantages, and this method has been used predominantly for the present paper. Many of the discharges mentioned here have since also been studied more carefully under stationary conditions.

2. EXPERIMENTAL PROCEDURE

A variety of discharge tubes were studied. The particular choice was chiefly due to what

[1] For a general review of gas discharges see, e.g., L. B. Loeb, *Fundamental Processes of Electrical Discharges in Gases* (John Wiley and Sons, Inc., New York, 1939); A. O. Engel and M. Steenbeck, *Elektrische Gasentladungen* (Edwards Brothers, Inc., Ann Arbor, 1944), 2 volumes; M. J. Druyvestijn and F. M. Penning, Rev. Mod. Phys. 12, 87 (1940).

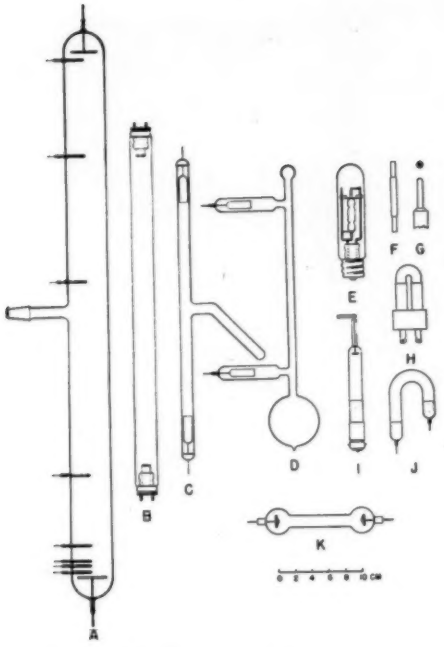

FIG. 1. Discharge tubes used in this study.
A. Air-discharge tube, variable pressure.
B. Low pressure mercury tube with argon (germicidal lamp).
C. Low pressure mercury tube with 2-mm argon (side tube to freeze out the mercury).
D. Tubes with Ne, A, K, X, at 2-mm pressure.
E. H–3 mercury lamp.
F. H–6 high pressure mercury lamp.
G. Low pressure capillary mercury tube.
H. High pressure mercury tube (Hanovia).
I. Mercury arc (no rare-gas filling).
J. Tube with mercury pool electrodes (no rare-gas filling).
K. Cadmium and zinc vapor arcs (with argon filling), Westinghouse.

was easily available. The gases with which the tubes were filled were the rare gases, air, mercury, cadmium, and zinc vapors, and certain mixtures.

Figure 1 gives a summary of the size and shape of the various tubes. The studies were made by observing the voltage across the tube, the current through it, and the intensity of the light as function of the time on the screen of an oscillograph. (Dumont Types 208B and 248.) The method of obtaining the electrical quantities is well known. The light intensity was obtained with the help of multiplier photo-tubes (R.C.A. Types 931A, IP21, or IP28) and the technique of doing this has been described previously.[2] The light intensity recorded was that of the integrated light at a definite spot in the tube or that of a particular spectrum line. The spot at which the intensity was observed could, if desired, be moved through the whole tube. The spectrograph used was usually a 21-foot grating in a Wadsworth mounting or a one-meter grating spectrograph.

Instead of observing voltage, current, or intensity as function of time, it is often more convenient to observe one of these quantities as function of one of the others. This is accomplished by feeding one signal to the horizontal, the other to the vertical deflection plates of the cathode-ray tube. Voltage *versus* current curves are called the characteristic of the discharge (example Fig. 2E). Examples of intensity *versus* current or intensity *versus* voltage are Figs. 2F and 2D.

In order to obtain a permanent record, photographs of the oscillograph screen were taken with a Leica camera on 35-mm film. The figures in this paper are reproductions of such photographs.

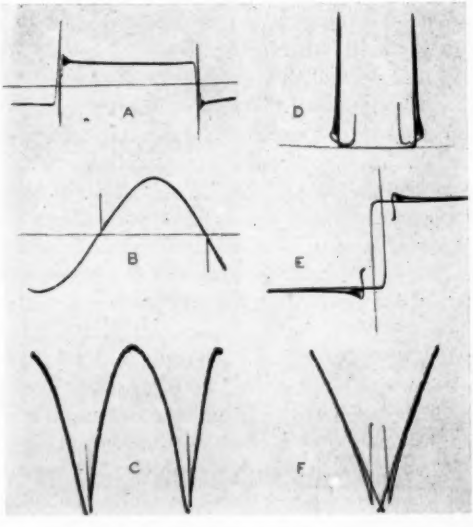

FIG. 2. Characteristics of typical low pressure mercury-glow discharge, (Fig. 1G) 60 cycles.
A. Voltage against time.
B. Current against time.
C. Intensity against time.
D. Intensity against voltage.
E. Voltage against current.
F. Intensity against current.

[2] G. H. Dieke, H. Y. Loh, and H. M. Crosswhite, J. Opt. Soc. Am. **36**, 185 (1946).

3. NORMAL GLOW DISCHARGE

Figure 2 gives the results for what we may call the normal case of a glow discharge. (Low pressure mercury tube Fig. 1G.) If we disregard for the time being the phenomena occurring at the beginning of each half-cycle, we see that the tube voltage remains practically constant during a half-cycle (Fig. 2A). The transformer produces a voltage which is nearly sine-shaped, and the current through the tube (2B) is roughly proportional to it as it is chiefly determined by the external resistance in series with the tube. The intensity of the light output (2C) resembles the current, except, of course, that it is rectified. That the light output is proportional to the current is more directly shown by the fact that the intensity *versus* current diagram (2F) is an approximately straight line. The characteristic curve (2E) is what we would get from 2A and 2B by eliminating the time. There is a phase difference between current and voltage expressed by the fact that ascending and descending vertical lines do not occur exactly at $I=0$. This is caused by the effect of the external circuit. Figure 2D again shows that the voltage remains constant during the time of light emission, which can also be deduced jointly from 2A and 2C.

This simple picture of the stationary state is modified at the beginning of each cycle by the short-lived breakdown phenomena. The voltage

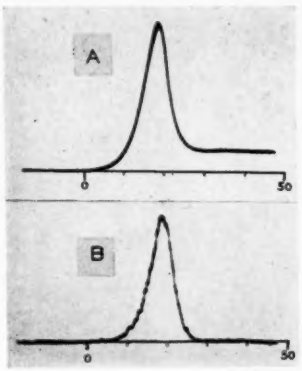

FIG. 4. Initial current and intensity surges at the breakdown of the discharge of Fig. 2. A. current; B. intensity; time in microseconds.

overshoots the stationary value often by more than 100 percent. When the breakdown occurs the voltage drops suddenly, usually below the stationary value, and there may be several oscillations about the stationary value with decreasing amplitudes before equilibrium is established (see e.g., Fig. 16B). The current at the time of breakdown has a very high instantaneous value, and may also execute several oscillations before taking on the value for the steady discharge. Figure 2B shows the breaks in the current for the beginning of the half-cycle, but the time resolution is too small to show details. The spark-like breakdown phenomena are much more accentuated when the frequency of the current is higher. Figure 3B shows clearly the current surges at the beginning of each half-cycle for an 800-cycle discharge, with the same tube. These breaks are so prominent at this frequency that the normal part of the discharge cannot be seen with the vertical scale of Fig. 3B. The current surge at the beginning of each half-cycle is accompanied by a surge in the intensity of the emitted light, which is shown as the sharp vertical line at the beginning of each half-cycle. (Figures 2C, 3E, F, and G.) It was shown pre-

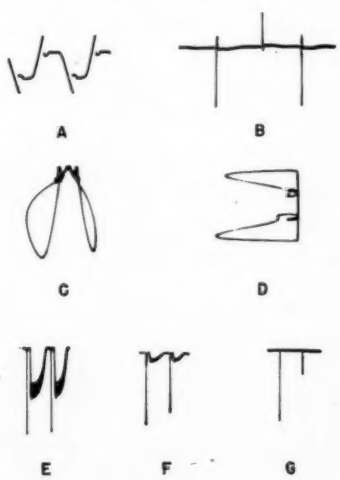

FIG. 3. Same tube as in Fig. 2, 800 cycles.
A. Voltage against time.
B. Current against time.
C. Intensity against current.
D. Voltage against intensity.
E. Intensity against time 2537A line.
F. Same 5461A line.
G. Same 3984A line of Hg II.

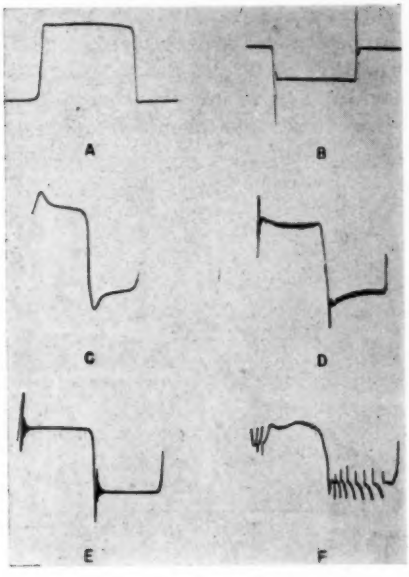

FIG. 5. Voltage diagrams of various discharges, 60 cycles.
A. Xenon tube 1D.
B. Mercury tube 1G, current 18 ma.
C. Mercury tube 1H.
D. Mercury tube 1B.
E. Mercury, tube 1G, current 4 ma.
F. Zinc vapor arc, tube 1K, current 4.3 amp.

viously[2] that this initial surge may be quite different for spectrum lines of different origin. The exact nature of the short-lived breakdown phenomena depends very much on the external circuit.

Figure 4 shows the details of the first breakdown on a very much extended time scale. The diagrams were triggered by the rise in voltage. They show that current and intensity surges are practically simultaneous, that the duration is of the order of a few microseconds, and that the intensity surge lasts slightly longer than the current surge.

An analysis of a discharge with the help of current, voltage, and intensity diagrams, like those shown in Figs. 2 and 3, can produce in a very simple way a great deal of significant information about the discharge. If a variety of discharges is investigated in this manner, we obtain a general idea of what type of phenomena are encountered. The tubes shown in Fig. 1 offer a good representation and they were all investigated in this way. The results obtained form a good foundation for further, more detailed studies.

In all these diagrams, mostly taken with 60-cycle alternating current, we distinguish, as in Fig. 2, the quasi-stationary part of the discharge and the breakdown. These parts may be sharply differentiated as in Fig. 5B or D, or there may be very little evidence of a breakdown as in Fig. 5A. The quasi-stationary part may show considerable deviation from that of a normal glow discharge, as in Fig. 5C or Fig. 7. Figure 6 shows the breakdown for a number of discharges in rare gases.

4. VARIATION WITH PRESSURE

The question immediately arises as to which physical parameters chiefly determine the character of a discharge. From even the most elementary observations we know that the gas pressure must be one of the most important variables. This is illustrated by Fig. 7 which shows a discharge in air (tube Fig. 1A) with the pressure ranging from atmospheric pressure down to the lowest at which a discharge still can be observed with an applied voltage of about 1000 volts. Figure 8 shows the current of the same tube. The figures show how strongly the pattern may be modified by changes in pressure.

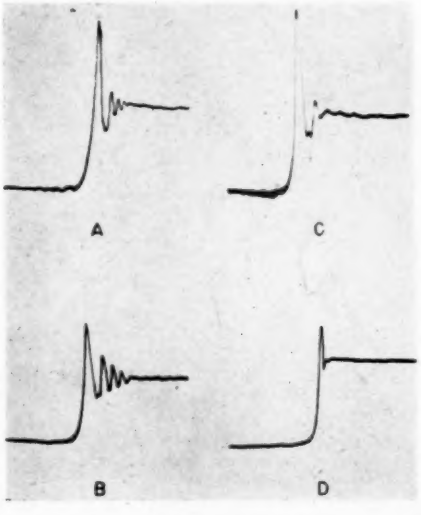

FIG. 6. Voltage oscillations in rare gases after breakdown, tubes 1D; A. Neon; B. Argon; C. krypton; D. Xenon.

5. MERCURY DISCHARGES

The influence of pressure can be observed easily in a mercury discharge, where the pressure in a sealed-off tube is a function of the temperature. Figure 9 shows the voltage pattern as a high pressure mercury tube warms up.

The mercury discharges, because they are easily available and important for many applications, were studied somewhat in detail. Figures 2–4 and Fig. 10 give details of a low pressure tube. Figure 11 presents the d.c. characteristic of this tube (Fig. 1G). The characteristic of the same tube with 60-cycle a.c. operation derived from the oscillograph traces, is given in Fig. 12. The d.c. characteristic is given for comparison. The difference in appearance of these curves in Figs. 11 and 12 is due to the difference in scale.

In the a.c. characteristic there is first the quasi-stationary part represented by the solid horizontal line at about 350 volts. The dotted part of the breakdown period deserves some comments. At the beginning of the half-cycle the voltage rises rapidly to nearly 800 volts without any current flowing. Then there is a sudden increase in current while the voltage drops. The suddenness of the phenomenon of course cannot be seen from Fig. 12 as the time has been eliminated. It can be seen from the current and voltage diagrams like Figs. 2A and B or qualitatively from the thickness of the traces on the characteristic which indicate how fast the cathode-ray beam passed over the curve. The figure shows that the tube voltage drops to quite large negative values. This, of course, can be caused only by the reaction of the external circuit. This part of the initial breakdown can be changed very considerably by changing the external circuit while the quasi-stationary part is largely independent of it.

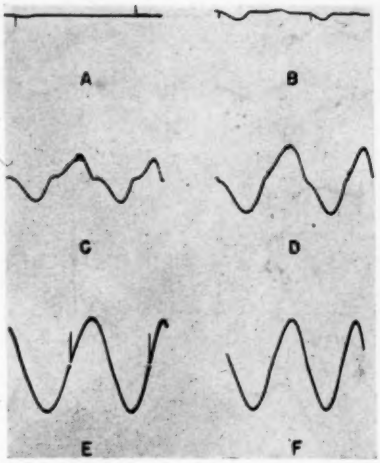

FIG. 8. Current diagrams of the same air discharges as Fig. 7.

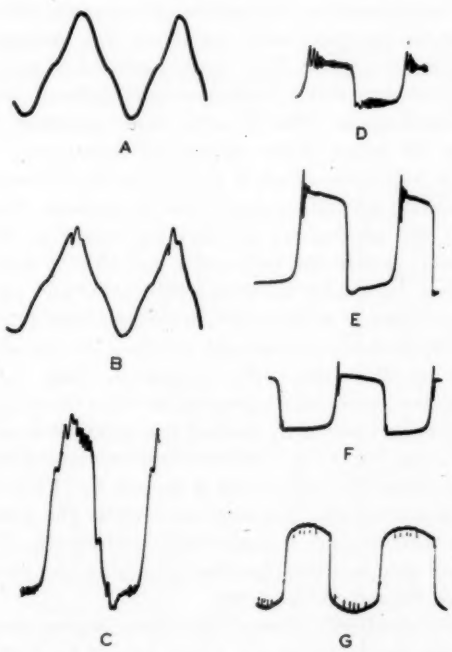

FIG. 7. Voltage diagrams of discharge in air with decreasing pressures. Tube Fig. 1A open circuit, D–F glow discharges.

Figure 12 shows that the quasi-stationary a.c. characteristic is not identical with the d.c. characteristic in this case. The reason for this is that for a given instantaneous value of the current the temperature of the discharge is different from that of the d.c. discharge at the same current. This means a different mercury pressure and, therefore, a different voltage.

This change in temperature may often be noticeable within one cycle, as will be apparent from some of the subsequent examples.

Figure 9 gives the voltage of a different type of mercury discharge where the length of the tube is sufficient to allow the development of a

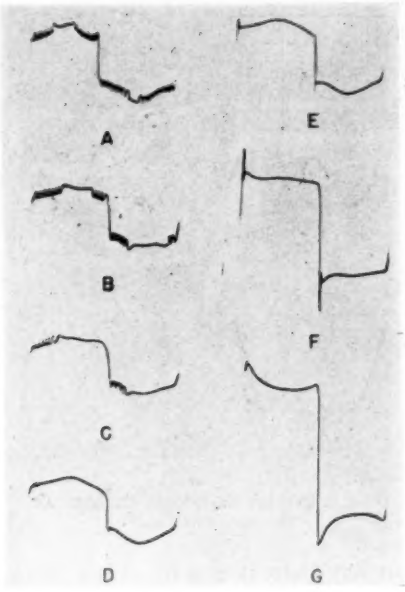

FIG. 9. Voltage pattern of a high pressure mercury-discharge tube during the warming-up period. Tube 1H. Time after starting: A–27 sec.; B–50 sec.; C–57 sec.; D–87 sec.; E–127 sec.; F–163 sec.; G–182 sec.

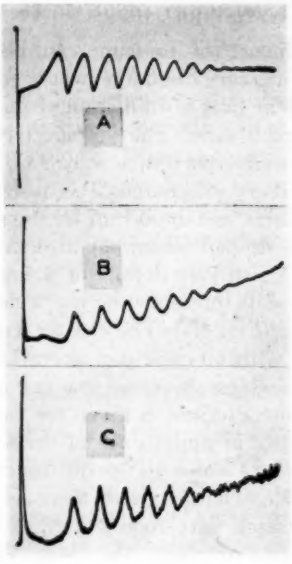

FIG. 10. Voltage, current, and intensity fluctuations accompanying the breakdown of discharge of Fig. 2. Total sweep about 800 microseconds.

positive column. When the tube is cold, the mercury pressure is very low and the discharge is started in the rare gas which is present in the tube at a pressure of a few millimeters. Within a few seconds, however, the mercury vapor takes over and the spectrum then contains only mercury lines, indicating that the rare gas does not play then any active part in the discharge. The mercury pressure gradually increases and the character of the discharge changes until equilibrium is established in about five minutes with a mercury pressure of the order of magnitude of one atmosphere. Figure 9 shows the change in voltage pattern; Fig. 13 that of the voltage-current characteristic as the tube warms up. Very typical are the voltage oscillations at the lower pressures which gradually disappear as the temperature is raised. These oscillations are not part or a consequence of the initial breakdown, but are related to the type of oscillations to be described later which are characteristic for so many types of discharges.

These oscillations are inhibited by the higher pressures. This is shown in Fig. 9A—where they are suppressed in the middle of the cycle where the temperature and, therefore, the mercury density is highest. They begin again when in the second half of the half-cycle the temperature is lowered again. This is even more apparent in Fig. 9B which shows clearly an asymmetry in each half-cycle which is due to the lag between temperature and current. After 57 seconds (Fig. 9C) the oscillations do not start again in the second part of the half-cycle, and this is again caused by the lag between temperature and current. After 87 seconds the oscillations have practically entirely disappeared and they are not visible at all at the higher pressures. They can, however, be made to appear again when the temperature is lowered by cooling the tube by an air blast (see Fig. 13I). That the characteristic of the tube cooled by an air blast is not *exactly* identical with one of the characteristics while the tube was warming up, is easily explained by the different temperature distribution within the tube while the outside is cooled.

The discharge current also shows fluctuations during the early stages. These cannot be easily seen in Fig. 13 because the scale is too small, but they are clear in Fig. 14 which shows the

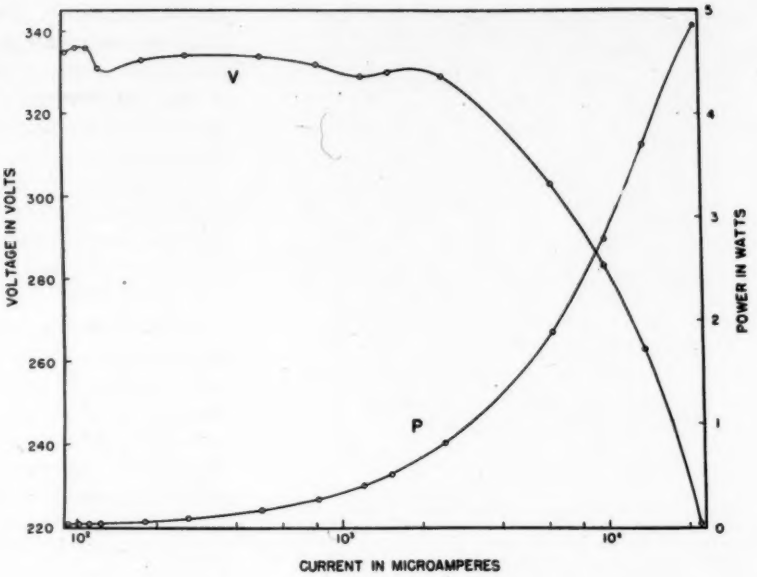

FIG. 11. d.c. characteristic of low pressure mercury discharge (tube 1G).

current pattern of this tube on an extended time scale.

The intensity as function of the time follows, in general, the current pattern, but the curve may be strongly affected by self-reversal. This is shown in Fig. 15 where simultaneous curves of the 2537A and 5461A mercury lines are obtained during the warming-up period of the same tube (Fig. 1H). An electronic switch served to alternate rapidly between the two lines, so that nearly continuous curves for both lines are obtained. Figure 15A shows how in the beginning the 2537 line is much stronger than the 5461 line, but after 30 seconds is strongly absorbed

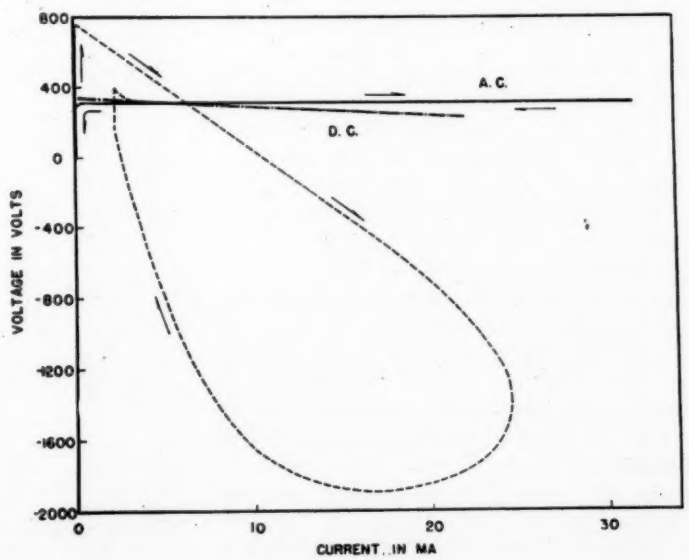

FIG. 12. Characteristic of low pressure mercury discharge (tube 1G).

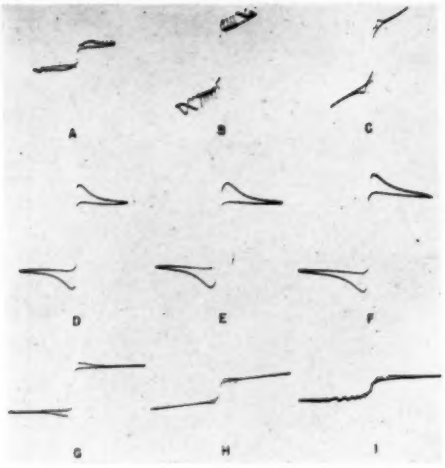

FIG. 13. Voltage-current characteristic of high pressure mercury discharge during warming up period. Time after starting: A–13 sec.; B–46 sec.; C–95 sec.; D–197 sec.; E–600 sec. Air blower turned on to cool the tube. Time after this F–22 sec.; G–38 sec.; H–54 sec.; I–83 sec.

by the vapor which has become denser. The shape of the 2537 curve has completely changed. Figures 15C and D show this state of affairs even more strongly. The green line in the meantime becomes very much stronger and, in the equilibrium state Fig. 15F, is many times stronger than 2537. In fact the lower trace is not due at all to the 2537 line proper, but to the continuous background in the vicinity of this line.

Figure 16 shows the effect of even higher pressure. The intensity of the 5461 line is appreciable even when the current and voltage are zero. This behavior is typical for sources at high pressure and may be due to two causes. When the current flow ceases there are a large number of ions present, and the light is radiated by recombination of the ions and step-wise transition of the lower states. Furthermore, the free electrons will still have an appreciable kinetic energy (high electron temperature) when the current ceases and, therefore, will be able to keep on exciting atoms. The relative contribution of the two processes must be left to more quantitative investigations of the future. The residual intensities are quite different for lines of different origin (see Figs. 16D, E, F). In general, the higher the excitation energy, the smaller the relative residual intensity. The fact that the stronger lines are more strongly self-reversed, and therefore are in reality much stronger than the peak of the intensity trace indicates, will also make the residual intensity appear stronger for such strong lines.

Experiments by Crosswhite on metal sparks have shown that the residual intensity depends very much on the excitation of the line.[3]

6. CADMIUM AND ZINC ARCS

Discharge tubes containing cadmium or zinc vapor are very similar in operation to the mercury discharges. We used such tubes obtained through Dr. H. C. Rentschler of the Research Laboratory of the Westinghouse Lamp Works. As most mercury tubes, they contain rare gases, and the discharge is started as a rare-gas discharge. As soon as the tube is heated by the discharge, enough metal is evaporated to carry the discharge. These two particular tubes were operated on 220 volts and drew 4–5-amperes current. Figure 17 gives the results of the cadmium tube. The zinc tube presents a similar behavior (see Fig. 5F). Figure 17A presents the voltage when the tube has come to equilibrium. It shows the typical over-shooting and break-

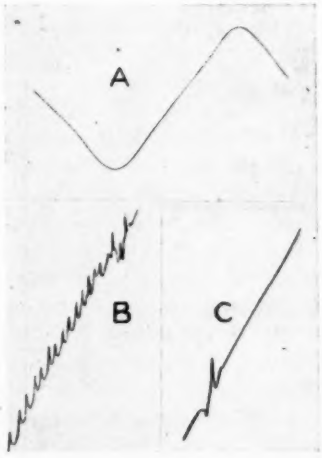

FIG. 14. Current fluctuations of early stages of same discharge as Figs. 9 and 13. A. complete cycle; B. details of the breaks; C. later stage.

[3] G. H. Dieke, "Symposium on Spectroscopic Light Sources, Buffalo, June 1946," Proc. A.S.T.M. (1947).

down at the beginning of each half-cycle, followed by a number of oscillations very similar to those seen in the mercury tube, represented in Figs. 9 and 13. The oscillations disappear in the second part of the half-cycle as the vapor pressure becomes too high because of the rise in temperature.

That the temperature rise is responsible for the disappearance of the oscillations is proven by Figs. 17B and C. In Fig. 17B the lower end of the tube was cooled by a blast of air, in Fig. 17C the upper end. We see that in either case the oscillations persist through one half-cycle. It appears that it is always the half-cycle during which the cooled end of the tube contains the anode. This shows that the oscillations must originate near the anode. As this is important for an understanding of their mechanism, proof of this fact is presented in Figs. 17D–F. Figure 17D shows the voltage pattern of the tube and a horizontal line obtained by applying zero voltage to the deflection plates. In Fig. 17E a constant positive voltage moves this line down.

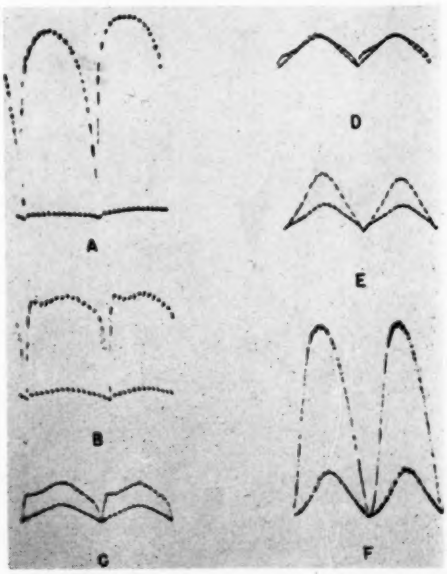

FIG. 15. Intensity pattern of the mercury lines 2537 and 5461 in high pressure tube 1G during the warming-up period. In A–D the stronger line is 2537, in E and F, 5461. Time after starting: A–15 sec.; B–30 sec.; C–120 sec.; D–170 sec.; E–195 sec.; F–300 sec.

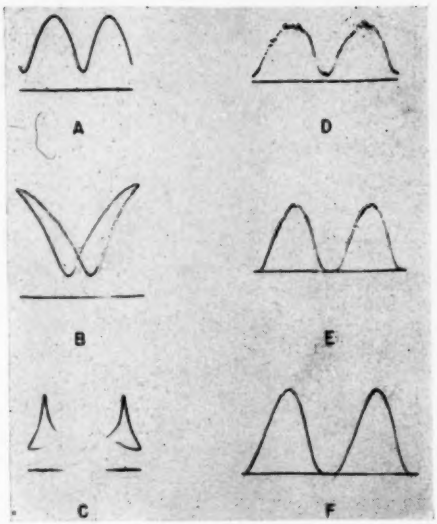

FIG. 16. High pressure mercury arc (H–3).
A. Intensity diagram of 5461 line.
B. Intensity against current.
C. Intensity against voltage.
D. 3131 line.
E. 3341 line.
F. Continuous background near 4000A.

In Fig. 17F it is apparent that the part of the tube which is positive is affected by the air blast.

7. INTENSITY OSCILLATIONS

Perhaps the most striking result revealed by the analysis of gas discharges with the cathode-ray oscillograph is the presence of very characteristic intensity oscillations. At first it was believed that such oscillations could only be observed in exceptional cases, but it soon became clear that they are a regular and characteristic feature of practically any gas discharge, provided the positive column is developed and the pressure is not too high.

Examples of oscillations at the beginning of each half-cycle were shown in Figs. 6 and 10. Figure 10 in particular shows that oscillations in voltage (A), current (B), and intensity (C) go together.

The oscillations to be described now have nothing to do with the breakdown. They are related to the so-called striations occasionally observed[4] in discharges more or less as curiosities.

[4] L. B. Loeb, reference 1, pp. 563, 573; Engel and Steenbeck, reference 1, p. 65.

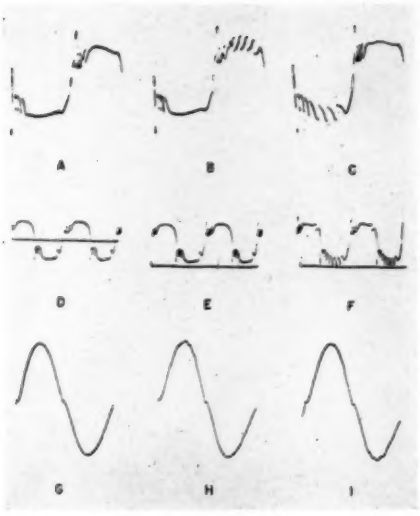

FIG. 17. Cadmium arc (tube Fig. 1K) current 4 amp.
A. Voltage diagram for equilibrium state.
B. Lower part of tube cooled by air blast.
C. Upper part of tube cooled by air blast.
D. Voltage with zero line.
E. Constant positive voltage added to same terminal that is connected to lower electrode.
F. Lower end of tube cooled by air blast. (The connections in D–F were reversed compared to those of A–C.)
G. Current diagram for equilibrium condition.
H. Lower end of tube cooled.
I. Upper end of tube cooled.

FIG. 18. Intensity breaks in a discharge corresponding to Fig. 9B or Fig. 13B.

Examples are found in the voltage fluctuations in the earlier stages of the discharge shown in Figs. 9 and 17. The voltage fluctuations are accompanied by current fluctuations (Fig. 14) and sudden breaks in the intensity as shown in Fig. 18. Such oscillations are also shown in some phases of the air discharge (Figs. 7 and 8).

However, much more typical cases are seen in Fig. 19 which shows the intensity pattern for a 60-cycle discharge in four tubes filled with neon, argon, krypton, and xenon, respectively, at approximately 2-mm pressure. (The intensities in this and some other figures are plotted downward.) Figures 19A–F show the intensity of the neon discharge at different places in the tube. Figure 19A is taken near one electrode. There are no oscillations when this electrode is the cathode, whereas they are quite pronounced in the half-cycle when this electrode is the anode. As soon as one gets away from the electrode there appear also oscillations in the cathode half-cycle and these oscillations are much more regular than those in the other half-cycle. This difference suggests the superposition of two different frequencies. The behavior for the other rare gases is similar, except that the frequency of the oscillations is different. This frequency is approximately 3360 c.p.s. for neon, 2760 c.p.s. for argon, 1800 c.p.s. for krypton, and 1200 c.p.s. for xenon.

Figure 20 shows a similar picture for the mercury tube, Fig. 1C, where again very typical in-

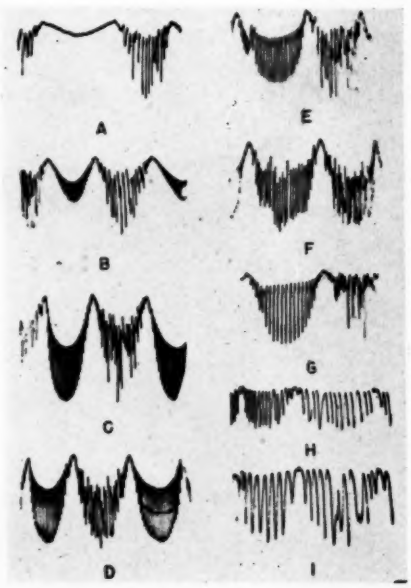

FIG. 19. Intensity oscillations in rare-gas discharges, pressure about 2 mm, 60-cycle discharge.
A. Neon, at one electrode.
B. Neon, in Crookes dark space for half-cycle with regular frequency.
C. Neon, negative glow.
D. Neon, Faraday dark space.
E. Neon, positive column.
F. Neon, near middle of tube.
G. Argon.
H. Krypton.
I. Xenon.

tensity oscillations occur. In the positive column very often the intensity is actually zero at the intensity minima. Further examples are given by Fig. 21 where A and B represent the current and voltage pattern of the discharge under the same conditions that produce the intensity pattern C. Both current and voltage look perfectly smooth and give no indication of the oscillations in the intensity. A closer inspection shows that there are actually also fluctuations in the voltage and current. Their amplitude, however, is only two or three percent of the total and, therefore, not visible on the scale of Figs. 21A and B, whereas the amplitude of the intensity oscillations is close to 100 percent.

Figures 21D, E, and F show similar fluctuation in an ordinary commercial low pressure mercury tube (so-called germicidal tube, Fig. 1B, which has the same type of discharge as the ordinary fluorescent tubes).

It is not difficult to ascertain experimentally that these oscillations are not oscillations pro-

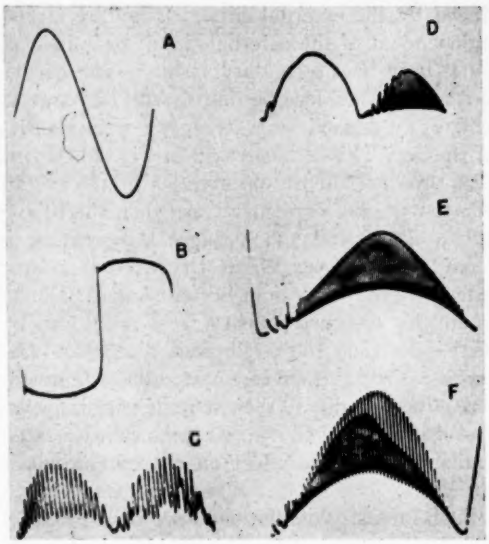

FIG. 21. A–current, B–voltage, C–intensity of low pressure mercury discharge, Fig. 1C. D–E, low pressure tube, Fig. 1B.

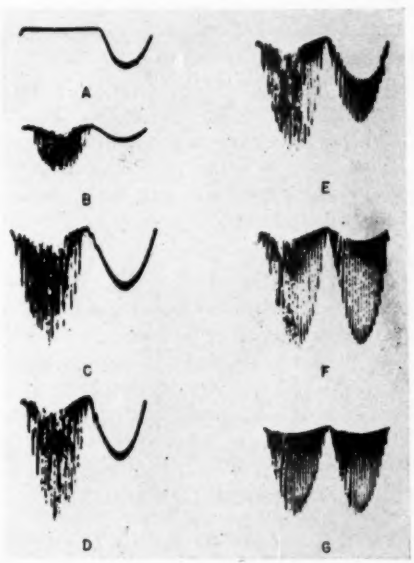

FIG. 20. Intensity oscillations in low pressure mercury tube; (Fig. 1C) current 30 ma, 60-cycle discharge.
A. At one electrode.
B. 0.4 cm from electrode, Crookes dark space.
C. 1.35 cm, negative glow.
D. 2.0 cm, Faraday dark space.
E. 4.4 cm.
F. 8.3 cm.
G. 12.8 cm, middle of tube.

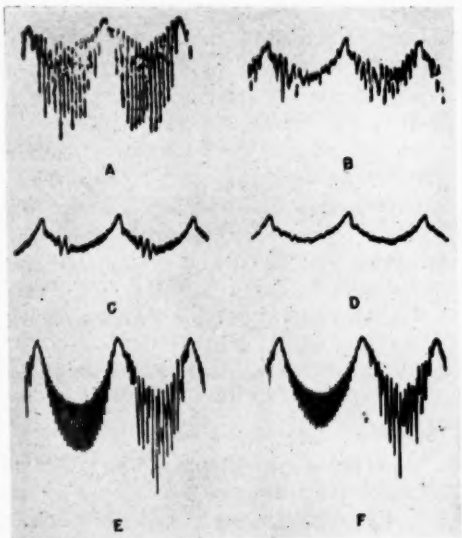

FIG. 22. A–D, low pressure mercury tube, Fig. 1D, intensity in middle of tube.
A. At thermal equilibrium, ambient temperature is room temperature.
B. After 2 minutes of heating with electric radiator.
C. After 3 minutes.
D. After 5 minutes.
Neon discharge.
E. Thermal equilibrium.
F. After 10 minutes heating by electric radiator.

duced by the external circuit. They are largely independent of the external circuit or influenced by it only in a secondary fashion. The oscillations must therefore be due to the discharge itself. They depend very strongly on the density of the gas. This is illustrated in Fig. 22. Figure 22A shows again the oscillations in the middle of a low pressure mercury tube when the tube is left to itself and the ambient temperature is room temperature. When the tube is heated externally the oscillations become less (22B), and when the mercury density is raised more by further heating, the oscillations disappear (22C and D). That this effect is not a direct temperature effect but due to the rise in density is shown by Figs. 22E and F, where a neon tube is heated similarly. Here the effect on the oscillations is negligible.

Such intensity oscillations have been found in practically every discharge that has been investigated. They form, therefore, not exceptional cases, but are part of the regular glow discharge. In the past the theoretical treatment of the glow discharge has usually dealt only with the stationary case. In the light of our experience it seems certain that the theory of the stationary discharge has been applied to many cases where the discharge is far from stationary. In fact, it seems that the fluctuating discharge represents the normal case of the glow discharge, and theory should account for the oscillations.

In order to clear up the mechanism of the oscillations, more observations under more carefully controlled conditions are needed. Such observations have been made in the meantime in our laboratory by Mr. T. Donahue, and have given valuable clues for a theoretical understanding of these phenomena. These observations will be reported in a subsequent paper.

Dr. H. M. Crosswhite has given considerable help in all stages of the preparation of this paper, which we gratefully acknowledge. We are greatly indebted to Dr. H. C. Rentschler for the construction of special tubes.

JOURNAL OF THE OTPICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Investigation of the Figure of Small Concave Metal-Replica Mirrors

JOHN N. SHIVE
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received July 16, 1947)

Optical tests were made on six paraboloidal metal electroplated replica mirrors of 3″ diameter and 2″ focal length, produced from convex glass masters. It was shown that the better replicas reproduced the features of the master surfaces with an average error of 5 or 6 microns per cm along the surface. Three of the replicas gave noticeably distorted images of a point source because of warping, introduced presumably in the process of separating them from the masters after electroplating.

INTRODUCTION

ELECTROPLATED metal-replica mirrors are used in applications where extreme perfection of figure is not required. They are made* by covering the surface of a master, usually of glass, with a thin coat of metal by chemical deposition, metalizing, vacuum evaporation, or sputtering. This metal film, still on the master surface, is then built up to a thickness of a few mm by electroplating, after which the replica so formed is mechanically separated from the master. Replicas are used for such purposes as searchlight reflectors, condensers for theater projectors, and heliograph mirrors.

During the war the development of radiometric devices for scanning a panorama by its own thermal radiation called for the inexpensive quantity production of small paraboloidal mirrors of a few inches diameter and aperture ratios of 1.0 or less. The replica process appeared to provide a ready solution for this need. It accordingly became necessary to learn with what degree of precision replicas could be obtained, how well replicas made from the same convex master resemble each other, and how well they reproduce the master.

GENERAL PROCEDURE

In an investigation undertaken to answer these questions, two convex paraboloidal glass masters, designated as Nos. I and II, were obtained. These were made by the U. S. Management and Engineering Company** under the direction of Dr. J. H. Hett. They were ground from standard 4″ Pyrex telescope blanks, of which only the central 3″ diameter portions were used in the optical tests to be described. The nominal focal length of each was 2″.

Two concave glass mates, similarly designated as Nos. I and II, respectively, were also made at the same time. The figures of these concaves were determined by zone-by-zone examination on an optical bench. The surfaces of the convex masters were then compared interferometrically with those of the corresponding concave mates. By adding the air film profiles thus determined to the figures of the concave mates, the figures of the convex masters were deduced.

The two masters were sent to the U. S. Army Engineer Board at Fort Belvoir, Virginia, and three 3″ concave metal replicas were made from each. Optical bench tests were subsequently made on these replicas. The three replicas made from each master were then inter-compared for similarity of figure, and their average figure was compared with the figure of the master.

DETERMINATION OF THE FIGURES OF THE TWO CONCAVE MATES

The concave mirror to be tested was placed in the holder of an optical bench and so adjusted that its axis passed through a test source. This consisted of a pair of flashlight lamps $1\frac{1}{16}$″ apart laterally and 45 feet away from the mirror. The image of this source formed by the mirror was projected onto a screen and examined with the aid of a low power binocular microscope. The distance between the two bright spots in the image served as a scale for estimating the diameters of the circles of least confusion. The screen was mounted on a carriage provided with a

* For details of this process see, for example, S. Wein, Metal Finishing **42**, 534 (1944).
** 401 Broadway, New York 13, New York.

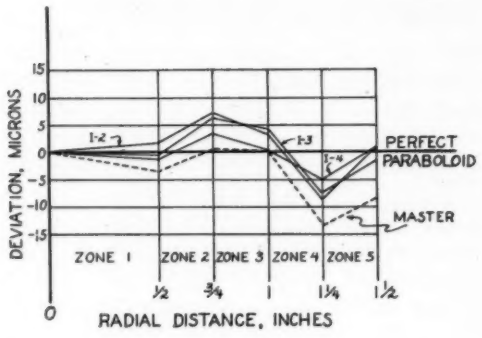

FIG. 1.

micrometer adjustment so that displacements of the screen along the axis could be measured to 0.001″. Five zone apertures were cut out of cardboard, of such radii that the entire surface of the mirror could be investigated zone by zone by interposing these apertures, one by one, into the optical path just in front of the screen. For each zone the diameter of the circle of least confusion and the screen-carriage micrometer setting at the focal position for least confusion were observed.

If the focal position of each of the several zones of a mirror is known, the approximate figure of its surface can be deduced from simple geometry. This determination of figure was made for each of the concave mates. The details need not be discussed here, but it may be of interest to state that for neither mirror was the departure from true paraboloidal figure over 10 microns at any place, nor was the gradient of departure from true figure anywhere more than 20 microns per cm along the radius of the mirror. The circle of least confusion at full 3″ aperture was about 0.3 mm for each mirror, while the resolving power of each was less than 0.1 mm in the image. The technical aspects of producing a paraboloidal surface 3″ in diameter with f:0.66 aperture having this image quality will be appreciated by any optical worker.

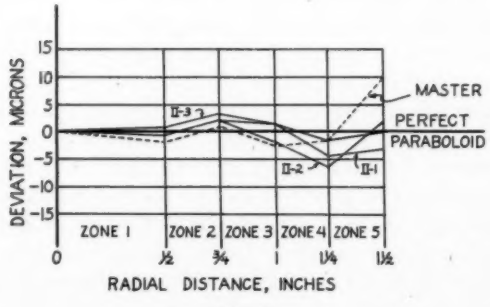

FIG. 2.

DETERMINATION OF THE FIGURES OF THE TWO CONVEX MASTERS

Each concave was placed in contact with its corresponding convex mate, and a study was made of the air film thickness gradients between the surfaces by the interference fringe method with sodium light. The predominating interference pattern in each case was one of radial symmetry with nearly circular fringes. From the number of fringes per unit radial distance the rate of change of air film thickness was determined at several radii for each pair. The direction of the thickness gradient at these various radii was determined by observing the direction of movement of the fringes as the eye was moved toward or away from the normal to the air film.

The figures of the convex masters were then obtained by adding the air film profiles thus determined to the figures of the concave mates.

EXAMINATION OF THE CONCAVE REPLICAS

When the six replicas were received they were examined by the same zonal focus and circle of confusion tests as have already been described for the concave glass mates. No care was taken to control the temperature during any of these tests, as it could be computed that a temperature variation by as much as 10°C would produce only insignificant changes in either the focal length of the replicas or the dimensions of the test equipment.

RESULTS AND DISCUSSION

The results of these various measurements on convex master No. *I* and the three replicas produced from it are presented in Fig. 1. A perfect paraboloidal figure is represented by the axis of abscissae, while deviations from this figure are given as ordinates. The origin refers to the center of the mirror. The figure of the convex master is given by the dotted curve, while the figures of the three No. *I* replicas, designated *I*-1, *I*-2,

and *I*-3, are shown by the three light solid curves, all referred to the same origin in plotting. The corresponding information for convex master No. *II* and its three replicas is presented in the same way in Fig. 2.

It is evident from an inspection of these plots that for the three replicas made from either master:

1. The profiles of the replicas do not differ from each other at any given radius by more than 5 or 6 microns, or from that of their master by more than 12 or 15 microns.
2. The radial gradients of the departure of the replicas from paraboloidal figure at any given radius do not differ from each other by more than 10 or 12 microns per cm, or from that of their master by more than 18 or 20 microns per cm.
3. The over-all average differences in profile and gradient are considerably less than the extremes just given.
4. The replicas correspond to each other somewhat better than they correspond to their master.

Replicas *I*-1, *I*-2, and *II*-1 exhibited evidences of warping, presumably introduced in the process of separating them from their masters after their formation. The various zones of these mirrors gave focal spots which were irregular stars or ellipses. The consequences of this warping do not show up in Figs. 1 and 2, for a warped zone in a replica still has about the same position of best focus which it would have if it were not warped: the image is merely poorer. The focal patterns of least confusion at full aperture for the warped mirrors were about 50 percent broader than for the undistorted ones. Such aberrations were also present without exception in a number of 6″, 7″, and 12″ replica mirrors from several sources examined by the author in 1944.

The results of this investigation show that, at its best, and when warping of the replicas is absent, the electroplated metal-replica process is capable of producing small mirrors which are quite satisfactory for most purposes not requiring extreme perfection of figure.

Catalog of the Adolph Lomb Optical Library at the University of Virginia

A catalog is now available which lists the publications contained in the Adolph Lomb Optical Library at the University of Virginia, of which Harry Clemons is Librarian. This very extensive collection will be of interest to workers in optics. Copies of the catalog, which has an introduction by James P. C. Southall, may be obtained by addressing the Librarian.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Focal Plane Diffraction Anomalies in Telescopic Systems*

A. I. MAHAN
Physical Optics Division, Naval Ordnance Laboratory, Washington, D. C.

(Received June 2, 1947)

A description is given of an anomaly in the form of the Fraunhofer diffraction pattern which is usually characteristic of a rectangular aperture. The anomaly is created by allowing two plane polarized monochromatic light waves of different amplitudes and phases, with their electric vectors polarized at various azimuth angles with respect to each other, to be incident on the upper and lower halves of the diffracting aperture and then combining these two waves in the usual way with a focusing lens. It was found that the forms and spacings of the diffraction bands parallel to the dividing line of the aperture could be changed quite markedly while those perpendicular to the dividing line were unaffected. The forms of these bands are calculated theoretically, and are also observed experimentally by a unique utilization of the Michelson interferometer. These experiments not only make it possible to change the resolving power of a telescope in directions perpendicular to each other, but also suggest the possibility of measuring linear distances in terms of changes in diffraction patterns instead of the usual shift of interference fringes. The forms of the diffraction bands in unpolarized light were also considered both theoretically and experimentally.

IN an earlier paper, the author in presenting a solution for the image doubling seen when looking through a telescope with an unusual roof prism mounted in front of the objective[1] found it necessary to solve a rather unique optical problem. The nature of the problem can be seen by referring to Fig. 1. Advancing upward from the plane of the figure were four accurately collimated plane electromagnetic light waves all of the same wave-length. The two waves which were restricted to the upper half of the aperture were polarized with their electric vectors perpendicular to each other and had amplitudes and phases which will be designated by A_1, φ_1, and A_3, φ_3. Similarly in the lower half of the aperture, there were two more plane polarized waves with their electric vectors parallel to the waves in the upper half of the aperture having amplitudes and phases which will be designated by A_2, φ_2, and A_4, φ_4. If now a telescope with an objective of sufficient aperture is placed over this complete region and these four emergent plane waves brought to a focus, a point image might be expected when looking into the eyepiece. However, when the dimensions $2a$ and $2b$ of the restricting aperture are made small enough and the power of the telescopic system large enough, horizontal and vertical sets of diffraction bands of the form shown in Fig. 2 appear in the focal plane of the telescope objective. As the amplitudes and phases of the four emergent waves are changed, the horizontal diffraction bands undergo marked changes in form while the vertical diffraction bands remain unchanged.[1]

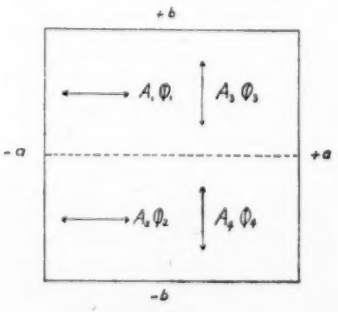

FIG. 1.

Actually before the above problem was solved,

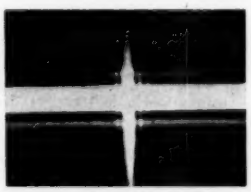

FIG. 2.

* Presented in part at the 31st Annual Meeting of Optical Society of America in New York City, October 3, 4, and 5, 1946.

[1] A. I. Mahan, J. Opt. Soc. Am. **34**, 623 (1945).

some simpler problems, which were partly discussed in the previous paper and which will be the subject of this paper, had also to be solved. Returning to Fig. 1, it should be obvious that one must first understand clearly the procedure for determining the forms of the two perpendicular sets of diffraction bands resulting from the focusing of two waves such as A_1, φ_1, and A_2, φ_2, or A_1, φ_1, and A_4, φ_4 before attempting a solution for the more complex problem of four waves which was discussed in the earlier paper.[1] Another closely related problem is that of determining the form of the horizontal and vertical diffraction patterns when the two waves in the upper and lower halves of the aperture are polarized at an arbitrary angle with respect to each other. Finally we will also be interested in the forms of the two sets of diffraction bands when the upper and lower halves of the aperture in Fig. 1 are illuminated with unpolarized light having different amplitudes and phases in the two regions. The present paper will discuss these four problems and will show a method whereby the theoretical forms of the diffraction patterns can be checked from photographs which were taken by utilizing a Michelson interferometer.

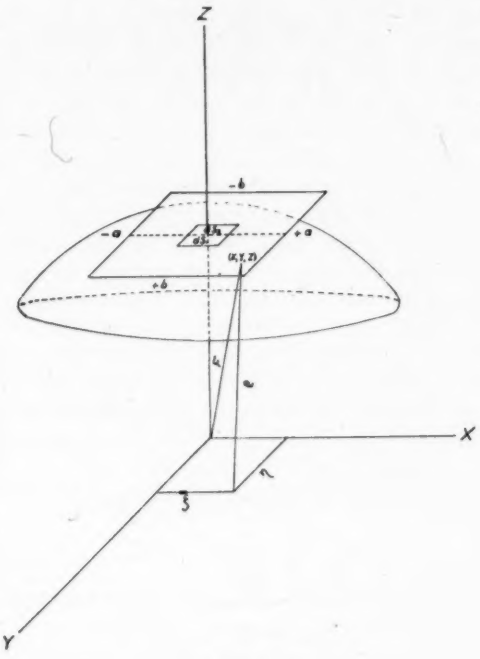

FIG. 3.

Apparently these problems are new ones. After solving them and searching the literature, however, it was found that G. B. Airy[2] in 1841 had solved a problem closely related to that of determining the forms of the horizontal set of diffraction bands for the A_1, φ_1, and A_2, φ_2 waves when he gave the first correct solution for the formation of Talbot's bands.[3] To observe Talbot's bands, all that is necessary is a white light source, a prism spectrometer, a small square restricting aperture and a thin piece of mica of uniform thickness (usually chosen by observing the interference colors in mercury light and using regions having the same color). The white light source is placed in front of the entrance slit of the collimator, and the viewing telescope with the square restricting aperture centrally placed over the objective is rotated until the continuous spectrum can be seen. If the collimator and telescope are properly focused for parallel light and the dimensions of the square restricting aperture and the thickness of the mica plate properly chosen, upon insertion of the mica plate over the left or right half of the aperture, the spectrum will be seen to be crossed by a series of usually rather faint parallel vertical bright and dark bands. In Airy's theory for the formation of these bands, he predicated that there would be present in the spectrum a series of diffraction bands of the form of the horizontal diffraction bands associated with the A_1, φ_1, and A_2, φ_2 waves later described in this paper. In observing Talbot's bands however, the detailed forms of these individual diffraction bands cannot be seen because of the overlapping of the diffraction bands from the many wave-lengths that are present. Thus although Airy never actually saw the diffraction bands which he postulated as being necessary for the formation of Talbot's bands, his theory nevertheless gave an accurate description of their forms as will be seen in this paper.

If one looks into the many text books on optics, it will be found that the subject of Talbot's

[2] G. B. Airy, Phil. Trans. Roy. Soc. London **130**, part 2, 225 (1840); **131**, part 1, 1 (1841).
[3] F. H. Talbot, London Edinburgh Phil. Mag. **10**, 364 (1837).

bands is usually avoided. In those books in which they are mentioned,[4] the reasons for their formation are very much simplified. After consulting Airy's original papers, the reasons for these approaches are evident. If the subject is to be treated in complete detail, it is too complex to be discussed in a text book. Unfortunately in these papers also, the more complex problem of the forms of the bands on both sides of the focal plane[5] is discussed first, while the simpler problem of the forms of the bands in the focal plane is looked upon as a special case of the more general problem of the band form on both sides of the focal plane. It will not be the purpose of this paper to go in detail into the theory of Talbot's bands but rather to make use of some of the ideas originally suggested by Airy which will be helpful to the understanding of the problems at hand. In this connection, therefore, the author would like to take this opportunity to present an independent more direct treatment of the problem of determining the form of the diffraction pattern resulting from the combination in the focal plane of the A_1, φ_1, and A_2, φ_2 waves. The resulting equation will be presented in a more complete form than that originally suggested by Airy and will also describe the form of the vertical diffraction bands which did not concern Airy in his problem. The theoretical forms of the diffraction bands in both the vertical and horizontal directions will be calculated for larger limits of the angular variable to show the changes in the secondary band structure which could not be seen in the case of the horizontal bands as predicted from Airy's original curves.

1. THEORY

The four waves in the upper and lower halves of the aperture shown in Fig. 1 will for simplicity be assumed to be sine waves. Considering first the case in which the two waves in the two halves of the aperture are polarized with their electric vectors parallel to A_1, φ_1, and A_2, φ_2 (or A_3, φ_3, and A_4, φ_4), these two waves at a time t as they emerge from the aperture can be written in the form

$$A_1 \sin(wt+\varphi_1), \quad A_2 \sin(wt+\varphi_2). \tag{1}$$

If now a lens (not shown in Fig. 3) is placed directly behind the diffracting aperture, these two waves after emerging from the lens might be expected to be transformed into a spherical wave with a radius of curvature of F (the effective focal length of the focusing lens) which would converge to a sharp focal point at the origin of the coordinate system shown in Fig. 3. When, however, the diffracting aperture is made small enough and the focal length of the telescopic objective long enough, some of the light will be diffracted to such a point as (ξ, η) in the focal plane $(x, y, 0)$ of the lens. The contributions of two very small wave surface elements dS_1 and dS_2 adjacent to each other on the z axis to the vibration in the focal plane at the point (ξ, η) can, according to Huygen's principle, be written in the form

$$d\bar{V} = KA_1\frac{dS_1}{\rho}\sin 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{\rho}{\lambda}\right)+KA_2\frac{dS_2}{\rho}\sin 2\pi\left(\nu t+\frac{\varphi_2}{2\pi}-\frac{\rho}{\lambda}\right). \tag{2}$$

In this equation ρ is the distance between the two small surface elements dS_1 and dS_2 and the point (ξ, η), λ is the wave-length and K a constant to take care of the physical dimensions of the equation. If now the length $2a$ and width $2b$ of the diffracting aperture are sufficiently small and the point (ξ, η) is not too far off the axis of the lens (z axis in Fig. 3), it is possible by utilizing the usual methods for Fraunhofer diffraction problems[6] to calculate the form of the diffraction pattern in the focal plane of the lens. The resultant vibration at the point (ξ, η) due to the two complete waves emerging from

[4] Arthur Schuster, *Introduction to the Theory of Optics* (Arnold Co., London, 1924) p. 120; R. W. Wood, *Physical Optics* (Macmillan Co., New York, 1934) p. 281.

[5] These bands in the strictest sense are not called Talbot's bands for this term is reserved only for those bands in the focal plane.

[6] R. W. Wood, *Physical Optics* (Macmillan Co., New York, 1914), p. 195.

the diffracting aperture of dimensions $2a$ and $2b$ can be written down directly. It is

$$\bar{V}=\frac{KA_1}{F}\int_0^{+b}\int_{-a}^{+a}\sin 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{\rho}{\lambda}\right)dxdy+\frac{KA_2}{F}\int_{-b}^{0}\int_{-a}^{+a}\sin 2\pi\left(\nu t+\frac{\varphi_2}{2\pi}-\frac{\rho}{\lambda}\right)dxdy. \tag{3}$$

The first term of the above equation represents the contribution of the A_1, φ_1 wave in the upper half of the aperture shown in Fig. 1 while the second gives the contribution of the A_2, φ_2 in the lower half of the aperture. In its present form, the above equation cannot be integrated. If Fig. 3 is again considered, ρ will for the more general case be found to be a known function of x and y. For an arbitrary point (x, y, z) in the region of the diffracting aperture, (see Fig. 3) it is quite easy to see that

$$\rho^2(x-\xi)^2+(y-\eta)^2+z^2 \tag{4}$$

and

$$F^2=x^2+y^2+z^2. \tag{5}$$

Substituting the above value for F^2 in Eq. (4), ρ will be found to have the value

$$\rho=\{F^2-2(x\xi+y\eta)+(\xi^2+\eta^2)\}^{\frac{1}{2}}. \tag{6}$$

If now a series development is made remembering that ξ and η are both small, this last expression reduces to

$$\rho=F-(x\xi+y\eta)/F. \tag{7}$$

Substituting this expression for ρ in Eq. (3), this equation now becomes integrable. It has the form

$$\bar{V}=\frac{KA_1}{F}\int_0^{+b}\int_{-a}^{+a}\sin 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{F}{\lambda}+\frac{x\xi+y\eta}{\lambda F}\right)dxdy$$
$$+\frac{KA_2}{F}\int_{-b}^{0}\int_{-a}^{+a}\sin 2\pi\left(\nu t+\frac{\varphi_2}{2\pi}-\frac{F}{\lambda}+\frac{x\xi+y\eta}{\lambda F}\right)dxdy. \tag{8}$$

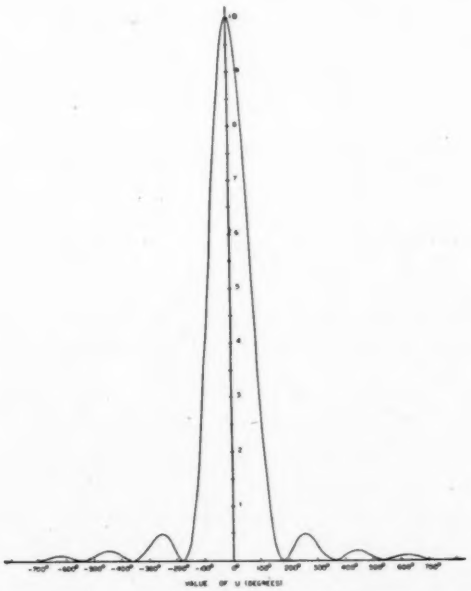

FIG. 4.

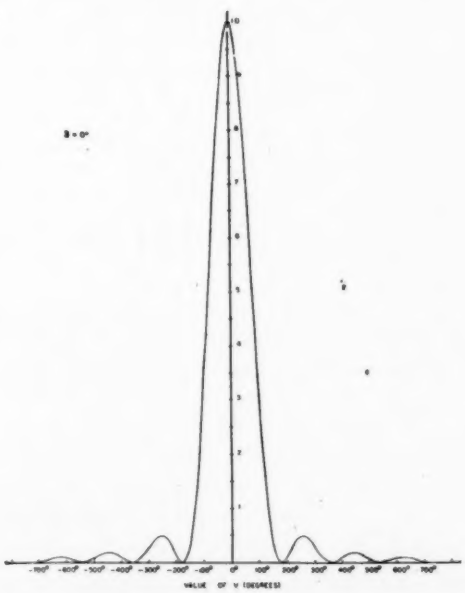

FIG. 5.

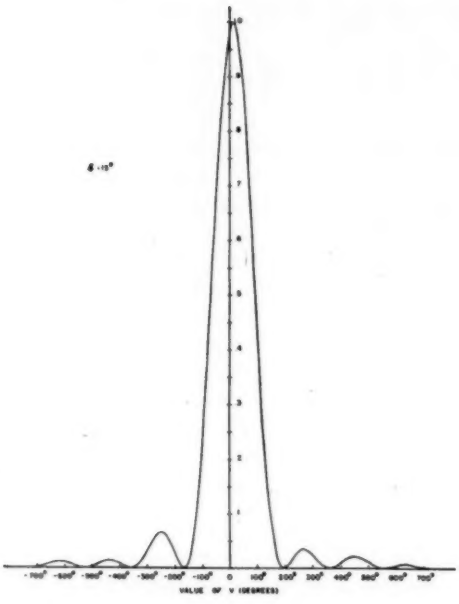

FIG. 6.

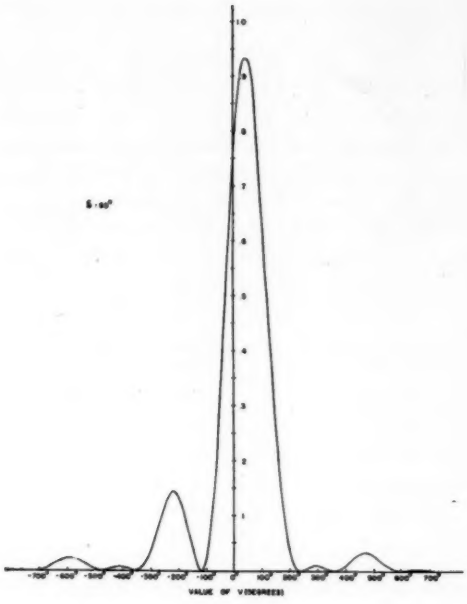

FIG. 7.

If each of these integrands is expanded completely grouping all terms not containing x and y in one term, this equation takes the following rather long form:

$$\begin{aligned}\bar{V}=&\frac{KA_1}{F}\sin 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{F}{\lambda}\right)\left[\int_{-a}^{+a}\cos\frac{2\pi x\xi}{\lambda F}dx\int_{0}^{+b}\cos\frac{2\pi y\eta}{\lambda F}dy-\int_{-a}^{+a}\sin\frac{2\pi x\xi}{\lambda F}dx\int_{0}^{+b}\sin\frac{2\pi y\eta}{\lambda F}dy\right]\\&+\frac{KA_1}{F}\cos 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{F}{\lambda}\right)\left[\int_{-a}^{+a}\sin\frac{2\pi x\xi}{\lambda F}dx\int_{0}^{+b}\cos\frac{2\pi y\eta}{\lambda F}dy+\int_{-a}^{+a}\cos\frac{2\pi x\xi}{\lambda F}dx\int_{0}^{+b}\sin\frac{2\pi y\eta}{\lambda F}dy\right]\\&+\frac{KA_2}{F}\sin 2\pi\left(\nu t+\frac{\varphi_2}{2\pi}-\frac{F}{\lambda}\right)\left[\int_{-a}^{+a}\cos\frac{2\pi x\xi}{\lambda F}dx\int_{-b}^{0}\cos\frac{2\pi y\eta}{\lambda F}dy-\int_{-a}^{+a}\sin\frac{2\pi x\xi}{\lambda F}dx\int_{-b}^{0}\sin\frac{2\pi y\eta}{\lambda F}dy\right]\\&+\frac{KA_2}{F}\cos 2\pi\left(\nu t+\frac{\varphi_2}{2\pi}-\frac{F}{\lambda}\right)\left[\int_{-a}^{+a}\sin\frac{2\pi x\xi}{\lambda F}dx\int_{-b}^{0}\cos\frac{2\pi y\eta}{\lambda F}dy+\int_{-a}^{+a}\cos\frac{2\pi x\xi}{\lambda F}dx\int_{-b}^{0}\sin\frac{2\pi y\eta}{\lambda F}dy\right].\end{aligned}\tag{9}$$

All the indicated integrations, however, in this expression are for simple sine or cosine terms and are, therefore, quite easily evaluated. When these integrals are evaluated and the common terms factored out, the previous equation reduces to

$$\begin{aligned}\bar{V}=&\frac{\lambda^2F^2}{2\pi^2\xi\eta}\cdot\frac{KA_1}{F}\cdot\sin\frac{2\pi a\xi}{\lambda F}\left[\sin\frac{2\pi b\eta}{\lambda F}\sin 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{F}{\lambda}\right)+\left(1-\cos\frac{2\pi b\eta}{\lambda F}\right)\cos 2\pi\left(\nu t+\frac{\varphi_1}{2\pi}-\frac{F}{\lambda}\right)\right]\\&+\frac{\lambda^2F^2}{2\pi^2\xi\eta}\cdot\frac{KA_2}{F}\cdot\sin\frac{2\pi a\xi}{\lambda F}\left[\sin\frac{2\pi b\eta}{\lambda F}\sin 2\pi\left(\nu t+\frac{\varphi_2}{2\pi}-\frac{F}{\lambda}\right)-\left(1-\cos\frac{2\pi b\eta}{\lambda F}\right)\cos 2\pi\left(\nu t+\frac{\varphi}{2\pi}-\frac{F}{\lambda}\right)\right].\end{aligned}\tag{10}$$

The two terms in each of the two component parts of this equation can be expressed in terms of a phase factory by multiplying and dividing by the square root of the sums of the squares of the co-

efficients of the time dependent since and cosine terms. If in addition the two substitutions

$$U=2\pi a\xi/\lambda F, \quad V=2\pi b\eta/\lambda F \tag{11}$$

are made and common terms factored out, the resulting expression becomes

$$\bar{V}=2ab\cdot\frac{K}{F}\cdot\frac{\sin U}{U}\cdot\frac{(2-2\cos V)^{\frac{1}{2}}}{V}\left\{A_1\sin\left[2\pi\left(\nu t-\frac{F}{\lambda}\right)+\varphi_1+\Delta\right]+A_2\sin\left[2\pi\left(\nu t-\frac{F}{\lambda}\right)+\varphi_2-\Delta\right]\right\}, \tag{12}$$

where

$$\tan\Delta=(1-\cos V)/\sin V. \tag{13}$$

To reduce this expression to a more usuable form, the two terms of this equation must be expanded. If in addition the radical is replaced in terms of half angles,

$$\bar{V}=2ab\cdot\frac{K}{F}\cdot\frac{\sin U}{U}\cdot\frac{\sin V/2}{V/2}\left\{A_1\sin 2\pi\left(\nu t-\frac{F}{\lambda}\right)\cos(\varphi_1+\Delta)+A_1\cos 2\pi\left(\nu t-\frac{F}{\lambda}\right)\sin(\varphi_1+\Delta)\right.$$
$$\left.+A_2\sin 2\pi\left(\nu t-\frac{F}{\lambda}\right)\cos(\varphi_2-\Delta)+A_2\cos 2\pi\left(\nu t-\frac{F}{\lambda}\right)\sin(\varphi_2-\Delta)\right\}. \tag{14}$$

Expanding the time independent sine and cosine terms still further and collecting terms,

$$\bar{V}=2ab\cdot\frac{K}{F}\cdot\frac{\sin U}{U}\cdot\frac{\sin V/2}{V/2}\left\{[A_1\cos\varphi_1\cos\Delta-A_1\sin\varphi_1\sin\Delta+A_2\cos\varphi_2\cos\Delta\right.$$
$$+A_2\sin\varphi_2\sin\Delta]\sin 2\pi\left(\nu t-\frac{F}{\lambda}\right)$$
$$\left.+[A_1\sin\varphi_1\cos\Delta+A_1\cos\varphi_1\sin\Delta+A_2\sin\varphi_2\cos\Delta-A_2\cos\varphi_2\sin\Delta]\cos 2\pi\left(\nu t-\frac{F}{\lambda}\right)\right\}. \tag{15}$$

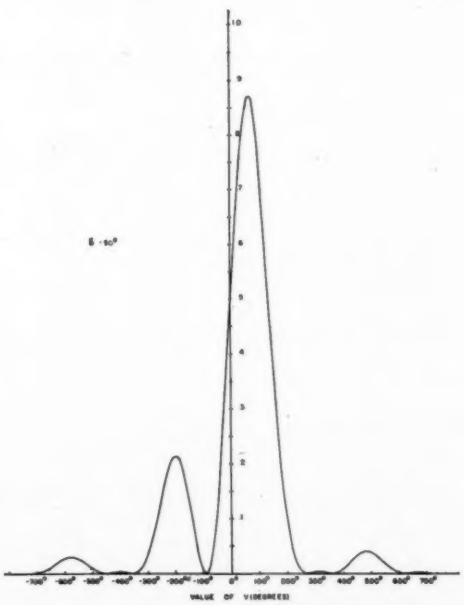

FIG. 8.

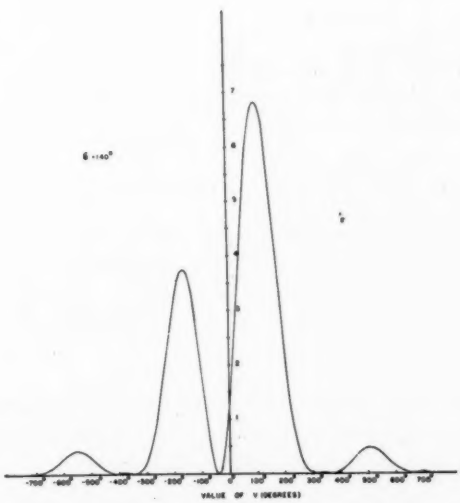

FIG. 9.

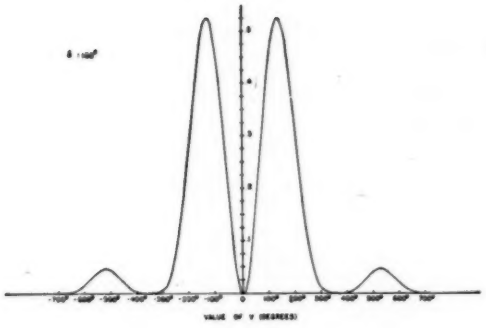

FIG. 10.

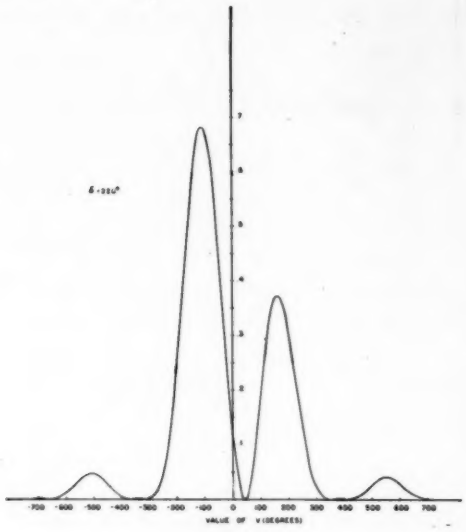

FIG. 11.

This equation can also be expressed in terms of another phase factor Φ by multiplying and dividing by the square root of the sums of the squares of the coefficients of the time dependent sine and cosine terms. Remembering that functions involving Δ can be evaluated in terms of V from Eq. (13), the resulting vibration in the focal plane of the telescopic objective has the final form given by

$$\bar{V}=2ab\cdot\frac{K}{F}\cdot\frac{\sin U}{U}\cdot\frac{\sin V/2}{V/2}\{A_1^2+A_2^2+2A_1A_2\cos(V-\delta)\}^{\frac{1}{2}}\sin\left[2\pi\left(\nu t-\frac{F}{\lambda}\right)+\Phi\right]. \tag{16}$$

In this equation δ is the difference in phase between the two waves in the two halves of the aperture. It, therefore, has the value

$$\delta=\varphi_1-\varphi_2. \tag{17}$$

The resulting intensity in the focal plane of the telescopic objective will be proportional to the square of the amplitude of the last expression or will be given by

$$J=\frac{(4ab)^2}{4}\cdot\frac{K^2}{F^2}\cdot\frac{\sin^2U}{U^2}\cdot\frac{\sin^2V/2}{(V/2)^2}\{A_1^2+A_2^2+2A_1A_2\cos(V-\delta)\}. \tag{18}$$

For the more simple case in which $A_1=A_2$ which is the one with which we shall be concerned experimentally, the intensity will be given by

$$J=\frac{(4ab)^2}{2}\cdot\frac{K^2A^2}{F^2}\cdot\frac{\sin^2U}{U^2}\cdot\frac{\sin^2V/2}{(V/2)^2}\{1+\cos(V-\delta)\}. \tag{19}$$

One last simplification can be made if the cosine term is expressed in terms of half angles. Making this substitution, the final form for the intensity will be given by

$$J=(4ab)^2\cdot\frac{K^2A^2}{F^2}\cdot\frac{\sin^2U}{U^2}\cdot\frac{\sin^2V/2}{(V/2)^2}\cdot\cos^2\frac{(V-\delta)}{2}. \tag{20}$$

The above equation predicts some rather interesting results. For instance, it says that the diffraction pattern in the U or ξ direction (parallel to the horizontal or x direction in Fig. 1) is independent of the phase difference δ and will be of the usual symmetrical form $\sin^2 U/U^2$ for rectangular apertures shown in Fig. 4. In the V or η direction, however (parallel to the vertical or y direction in Fig. 1), the form of the diffraction pattern will be a function of δ.

This dependence of the form of the diffraction pattern in the η direction on δ is shown in Figs. 5 to 12 for the values of δ indicated. For purposes of plotting evaluations were made at every 10° along the V scale from $-720°$ to $+720°$. At $\delta=0°$, (see Fig. 5) the pattern as might be expected is of the same symmetrical form as the diffraction pattern in the ξ (or U) direction shown in Fig. 4. For small increases in δ, however, the chief maximum moves to the right and decreases slightly in intensity (see Fig. 6). The first maximum on the left side simultaneously begins to increase in intensity and also move to the right while the first maximum on the right decreases in intensity and moves to the right. With further increases in δ, the chief maximum continues to move to the right and decrease in intensity while the first smaller maximum on the left increases and continues to move to the right, and the first small maximum on the right decreases in intensity and moves to the right (see Figs. 7, 8, and 9). At $\delta=180°$ (see Fig. 10), the first small maximum on the left has grown in intensity until it is of the same intensity as the former chief maximum while the first small maximum on the right has disappeared. It will be noted also that the positions of the maxima of these two equally intense bands are located just above the positions of the first minima for $\delta=0°$ (see Figs. 5 and 10). Beyond $\delta=180°$, the maximum on the right continues to decrease in intensity and move to the right while the maximum on the left increases simultaneously in intensity and moves to the right to make room for a smaller band which appears on its left side increases in intensity and also moves to the right. At $\delta=360°$, the original chief maximum has decreased to the same size and has moved into the position of the first small maximum on the right of the central position; the original small maximum on the left has increased to the same size and has moved into the central band position; the new band appearing to the left of the strong maximum in Fig. 12 has increased to the same size and has moved into the position of the first small maximum on the left of the central band shown in Fig. 5. For each change in δ of 360°, the cycle is always such that the pattern is repeated and each band in the central band structure ($-360°$ to $+360°$) has moved to the right one position. If the patterns are evaluated for decreasing values of δ, the changes can be made to take place in reverse order and the fringes will move to the left.

Changes in the smaller outside band structure also appear as δ changes.[7] For $\delta=0°$, (see Fig. 5) it will be noted that for values of V on each side of the axis between 360° and 720° two fringes are present. As δ increases, the inner fringe on the right hand side of the central band structure between these limiting values of V increases in intensity, broadens and moves to the right, the outer fringe simultaneously decreasing and moving to the right until at $\delta=180°$, the inner fringe occupies the same space formerly held by the two smaller fringes. On the left-hand side of the central band structure, the outer smaller fringe between these same limiting values of V increases in intensity with increasing values of δ, broadens, and moves to the right while simultaneously the larger fringe on the right decreases in intensity and moves to the right until at $\delta=180°$ again the fringe on the left occupies the same space formerly held by the two smaller fringes. For values of δ beyond 180° (see Figs. 11 and 12), the maximum on the right of the central band structure continues to move to the right but now decreases in intensity and becomes narrower to make room for a new band on its left side which simultaneously increases in intensity, broadens, and also moves to the right. The corresponding maximum to the left side of the central band structure (see Fig. 10) also continues to move to the right and decreases in intensity and becomes narrower to make room for the new band which appears on its left side and increases in intensity, broadens and also moves to the right. When δ has gone through a change of 360°, the cycle is always such that the original pattern is repeated and each

[7] These changes in the outer band structure could not be seen in Airy's original curves because the limiting values of V extended from $-360°$ to $+360°$.

band in the outer band structure has moved to the right by one position. If patterns are evaluated for decreasing values of δ instead of increasing values, the movement of the fringes will be to the left.

If now we consider the waves A_1, φ_1, and A_4, φ_4, (or A_3, φ_3, and A_2, φ_2) which are polarized perpendicularly to each other, the problem is very similar. Returning again to Fig. 1, it is quite evident that the limits for the integration of the A_4, φ_4 wave are identical with those for the A_2, φ_2 wave. As might be expected, therefore, the equations up to and including Eq. (12) from the previous problem will be applicable here with the A_2 and φ_2 replaced by A_4 and φ_4. Beyond this point, however, the treatment of the present problem is simpler. Rewriting Eq. (12) so that it applies to the A_1, φ_1, and A_4, φ_4 waves, and using the suggested substitution for half angles, the resultant vibration in the focal plane is

$$\bar{V}' = 2ab \cdot \frac{K}{F} \cdot \frac{\sin U}{U} \cdot \frac{\sin V/2}{V/2} \left\{ A_1 \sin\left[2\pi\left(\nu t - \frac{F}{\lambda} \right) + \varphi_1 + \Delta \right] + A_4 \sin\left[2\pi\left(\nu t - \frac{F}{\lambda} \right) + \varphi_4 - \Delta \right] \right\}. \tag{21}$$

The two component parts of the bracketed expression are now made up of two waves which are polarized with their electric vectors vibrating perpendicular to each other, and, therefore, cannot interfere with each other. The resultant intensity in the focal plane will then be proportional to the sums of the squares of the amplitudes of these two expressions or may be written

$$J' = (4ab)^2/4 \cdot K^2/F^2 \cdot \sin^2 U/U^2 \cdot \sin^2(V/2)/(V/2)^2 \cdot (A_1^2 + A_4^2). \tag{22}$$

If again $A_1 = A_4$ which is the case with which we shall be concerned experimentally, this expression reduces to a final form given by

$$J' = (4ab)^2 \cdot \frac{K^2 A^2}{F^2} \cdot \frac{\sin^2 U}{U^2} \cdot \frac{\sin^2 V/2}{2(V/2)^2}. \tag{23}$$

This equation is quite similar in form to Eq. (20) with the exception that the cosine term involving δ is missing. Since δ is absent from this equation, the diffraction patterns in both the vertical and horizontal directions are independent of δ. In the U or ξ direction (parallel to the horizontal or x direction in Fig. 1), the diffraction pattern is of the same form as that for the previous problem which was shown earlier in Fig. 4. In the V or η direction, however, (parallel to the vertical or y direction in Fig. 1), the form of the pattern is given by $\frac{1}{2}[(\sin^2(V/2)/(V/2)^2]$. It, therefore, has the same symmetrical form but the bands are twice as wide as those in Fig. 4 (see Fig. 13). In this direction, the diffraction pattern is of the same form as that characteristic of an aperture half the height of Fig. 1.

The previous two problems immediately suggest a third problem namely, what is the form of the diffraction pattern if two collimated monochromatic plane waves in the upper and lower halves of the aperture are polarized at an arbitrary angle α with respect to each other. If the amplitudes of these waves are designated by A_u and A_d, it is possible from Eqs. (18) and (22) by separating out the parallel and perpendicular components to write down immediately the intensity of the diffracted light. It is

$$J'' = \frac{(4ab)^2}{4} \cdot \frac{K^2}{F^2} \cdot \frac{\sin^2 U}{U^2} \cdot \frac{\sin^2 V/2}{(V/2)^2} \{ A_u^2 + A_d^2 + 2A_u A_d \cos\alpha \cos(V - \delta) + A_d^2 \sin^2\alpha \}. \tag{24}$$

Remembering that $\sin^2\alpha + \cos^2\alpha = 1$,

$$J'' = \frac{(4ab)^2}{4} \cdot \frac{K^2}{F^2} \cdot \frac{\sin^2 U}{U^2} \cdot \frac{\sin^2 V/2}{(V/2)^2} \{ A_u^2 + A_d^2 + 2A_u A_d \cos\alpha \cos(V - \delta) \}. \tag{25}$$

If $A_u = A_d$,

$$J'' = (4ab)^2 \frac{K^2 A^2}{F^2} \cdot \frac{\sin^2 U}{U^2} \cdot \frac{\sin^2 V/2}{2(V/2)^2} \{ 1 + \cos\alpha \cos(V - \delta) \}. \tag{26}$$

As a check on this equation, it should reduce to Eq. (19) when $\alpha = 0°$ and to Eq. (23) when $\alpha = 90°$ and such will be found to be the case. Since no attempt was made experimentally to check this equation, it will simply be left in this final form without further discussion.

The last problem which should be discussed is that of the form of the diffraction pattern to be expected when both the upper and lower halves of the aperture are illuminated with collimated, monochromatic, plane, unpolarized light waves having different amplitudes and phases in the two regions. If one assumes the amplitudes of the electric vectors and the phases of the two waves in the upper and lower halves of the aperture to be given again by A_1, φ_1, and A_2, φ_2, for each A_1, φ_1 wave entering the diffracting aperture at a given azimuth angle, there will be a corresponding wave A_2, φ_2 in the lower half of the aperture with its electric vector parallel to that of the upper half of the aperture. To obtain the diffraction pattern for any given azimuth angle and a given value of δ, it is necessary to use Eq. (18) again. Since unpolarized light is usually regarded as having all possible azimuth angles present, the resultant diffraction pattern for a given value of δ could then be obtained directly by summing Eq. (18) over all possible azimuth angles. Without writing it down, therefore, the diffraction pattern for unpolarized light with any given value of δ would be expected to be of a different intensity but to have the same form as that predicted by either Eq. (18) or (20).

2. EXPERIMENTAL VERIFICATION

To set up experimentally a system which will check the forms of the diffraction patterns for the conditions under which they were derived theoretically would be impossible. For instance in the first three cases which were discussed, two plane polarized light waves of fixed amplitude and phase were assumed to be entering the diffracting aperture. Experimentally such waves are impossible to obtain, for ordinary gaseous light sources even in the simpler classical sense are visualized as having many oscillators vibrating with varying amplitudes and different phases. If, at any instant, one collimates such a group of waves and allows the resulting plane waves to pass through a polarizing Nicol prism, all components of the electric vectors vibrating along the direction of transmission of the Nicol prism are allowed to pass while those components vibrating perpendicularly are not transmitted. At a given instant the emerging plane polarized waves which are transmitted still have a wide range of amplitudes and phases. For unpolarized light similarly the assumptions of constant amplitude and phase for identical azimuth angles in the upper and lower halves of the aperture are also not realizable experimentally. Carefully collimated monochromatic light has also been assumed. It might seem, therefore, that the theory is not susceptible to checking by experimental methods.

Actually if one looks a little at Eqs. (20) and (23) which are the only two with which we shall be concerned experimentally, some interesting facts appear. In Eq. (20) for instance if one does have two plane waves of the same amplitude polarized in parallel directions to each other in the upper and lower halves of the diffracting aperture (see Fig. 1), for a given geometry of diffracting aperture and focusing lens, the dif-

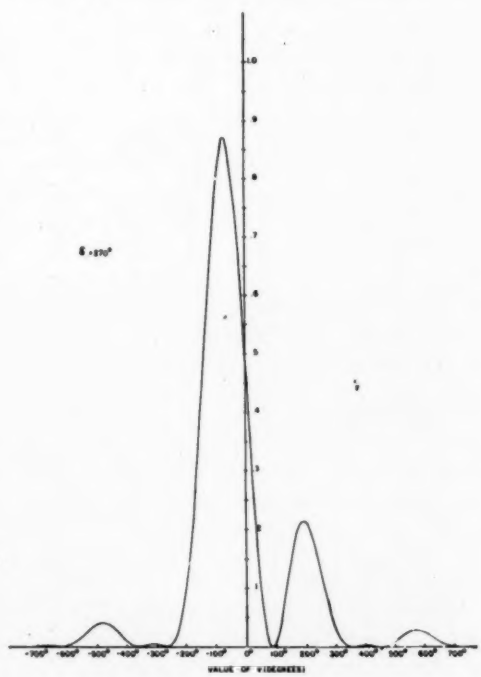

FIG. 12.

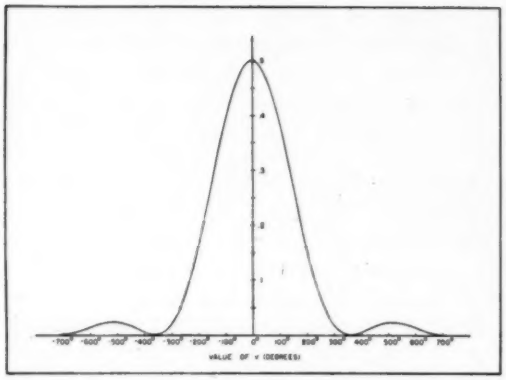

FIG. 13.

fraction pattern form depends then only on the difference in phase between the two waves in the two regions of the aperture. If, instead of a single wave of constant amplitude in each half of the aperture, one has many waves of varying amplitudes and different phases entering at the same azimuth angle and so arranged that each wave in the upper half of the aperture has a corresponding wave in the lower half with the same amplitude but always having the same constant difference in phase, the form of the diffraction bands formed by such an array of waves can be arrived at in two ways. At any instant the two corresponding waves in the two regions of the aperture can be combined separately as in Eqs. (16) and (21) and all these resultant focal plane displacements combined into a single displacement from which the intensity can be evaluated from an equation of the same form as Eq. (20). A second approach is that of combining at any given instant all the waves in the two regions of the aperture into a resultant wave and then evaluating the diffraction pattern form directly from Eq. (20). Whichever method is used the resulting diffraction pattern form should be the same. In the case of Eq. (23), the form of the diffraction pattern as might be expected is independent of the phase difference δ since it does not appear in the equation. The only requirement from this equation for observing the diffraction bands of the form calculated is that both halves of the aperture shall be illuminated by waves having the same amplitude. This is not a necessary requirement for the diffraction bands will still be of the same form even though the amplitudes are different as is indicated by Eq. (22). In the presence of many waves of different amplitudes and phases as before the waves in each of the two halves of the aperture may be combined in pairs in the focal plane or all the waves in each half combined separately and the two resultant waves then combined in the focal plane to obtain the final diffraction pattern. The resultant diffraction pattern which is viewed or photographed will in each case then theoretically be the time average of all these independently calculated diffraction patterns. Since at any instant these calculated patterns will be of the same shape but will differ only in magnitude, the resulting time average will be of the same general form as that calculated at any time for any known set of conditions.

According to the discussion of the previous paragraph, if the plane waves in the upper and lower halves of the aperture in Fig. 1 are polarized with their electric vectors parallel to each other, they must be coherent waves if the type of diffraction bands described by Eq. (20) are to be observed. When the electric vectors in the two regions of the aperture are polarized perpendicularly to each other, the waves may either be coherent or incoherent and the diffraction pattern described by Eq. (23) will still be observed. The instrument which immediately suggests itself for the purpose of obtaining two coherent plane waves is the Michelson interferometer. This instrument in its usual form with carefully collimated light makes it possible to secure two coherent plane waves exactly superimposed on each other. The apparatus which was finally used for recording the diffraction pattern forms is shown in Fig. 14. A cross consisting of lines about 0.01 mm wide was ruled on a silvered glass plate and mounted behind the hole in the large box. An AH4 high pressure mercury arc, green filter for the 5461A line, and an auxiliary lens were also placed inside the box in such a position as to illuminate the previously mentioned cross. This box then served to provide a monochromatically illuminated target for all photographs, the room being completely darkened for all exposures. A 500-mm E.F. telescopic objective was mounted in front of the illuminated cross at a distance equal to its back focus so

that parallel light emerged from the opposite side. This parallel beam of light was then allowed to fall on the first surface of the dividing plate of a Michelson interferometer. At this point the light is divided into two beams one traveling along one of the interferometer arms and the second along the arm at right angles to the first. If the reflecting surface of the beam-splitting plate is coated with a high reflecting film to utilize the correct set of fringes, it is quite easy by considering reflection and absorption losses to show that the two beams emerging from the interferometer are nearly of the same intensity. The light after emerging from the interferometer was incident upon a centrally located 5-mm square aperture which was inserted to magnify the diffraction effects. This aperture was made by gluing four razor blades on the white cardboard which is mounted just in front of the interferometer in Fig. 14. A 2839-mm long focal length lens was mounted directly behind this diffracting aperture to image the initial cross with a magnification of about six times. The interferometer was so adjusted that the optical paths were the same (white light fringes in the field of view) and the mirrors oriented so that the complete field was filled by a single fringe. This adjustment makes it certain that the phase difference between the two waves over the complete region is constant which is necessary if our initial conditions in Fig. 1 are to be met. If the optical paths are made too much different, Haidinger fringes will in general begin to appear because of the angular spread of the beam from the collimator. Theoretically these may be made to disappear by reducing the source to a point source, but in doing so the intensity of the diffraction bands are tremendously reduced making them very difficult to photograph. To simulate further the conditions for the problem as suggested by Fig. 1, a razor blade fastened to the auxiliary interferometer plateholder was mounted in the lower half of the interferometer arm at the right. This meant that light from the upper half of this interferometer arm passed through the upper half of the 5-mm square aperture. Similarly a second razor blade fastened to the tripod stand was placed in the upper half of the rear interferometer arm so that light from the lower half of this interferometer arm entered the lower half of the 5-mm square aperture. Light in the lower half of the aperture, therefore, comes from the rear interferometer arm, while light in the upper half of the aperture comes from the right interferometer arm. The phase difference between the unpolarized waves in the two regions of the aperture can be changed at will by simply moving the mirror in the rear interferometer arm providing, of course, that this displacement is not too far away from the position of zero path difference so that the central Haidinger fringe in the absence of both razor blades in the arms of the interferometer no longer fills the 5-mm square aperture. To check the phase difference between the two unpolarized waves in the two regions of the aperture, all that is necessary is to slide the razor blade in the rear interferometer arm out of the field of view. If the upper half of the field of view looking through the 5-mm square aperture is dark, the two waves are out of phase; if it appears bright, they are in phase.

To further fulfill the conditions for the case of polarized light suggested by Fig. 1 and Eqs. (20) and (23), it is necessary to have some means of polarizing the light in different directions in the upper and lower halves of the field. In principle this might readily be done by simply inserting polaroid filters over the upper and lower halves of the 5-mm square aperture cut so that the planes of polarization are at any desired angles with respect to each other. In practice this is not very easy for it must be remembered

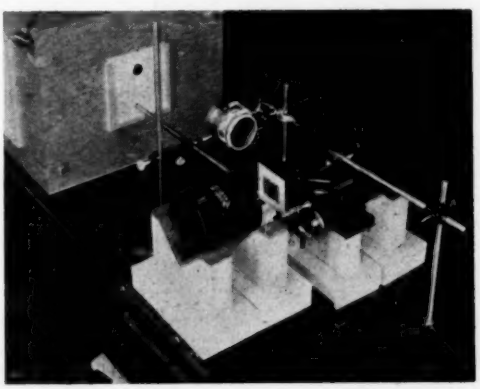

FIG. 14.

that the difference in phase between these two waves must be constant over all regions of both apertures. This means that such filters ought to be of interferometer quality and usually they are far from it. In the case in which both waves are polarized in parallel directions, this difficulty may be avoided by placing the polarizer between the source and the initial illuminated cross. In the case in which the two waves are polarized in perpendicular directions, one might be fortunate enough to obtain a good half-wave plate which can be placed over either the upper or lower half of the 5-mm square aperture. Such a wave plate ought to show uniform color in monochromatic light.

The following adjustments were made to obtain the diffraction pictures. The large focusing lens and the razor blade from the rear interferometer arm were removed from their respective places. Assuming that the center of the illuminated cross has been placed on the axis of the collimating lens and that this axis lies in a plane through the central point of each interferometer plate and mirror and the 5-mm aperture, by rotating the interferometer about a vertical axis the light beam coming out from the interferometer can be made to emerge at an angle of 45° with respect to the last reflecting surface. The interferometer is next adjusted so that the two arms are equal and the central fringe fills the complete field of view. The micrometer screw for moving the rear mirror is then rotated slowly and the number of revolutions determined to the nearest fraction between two points at which the upper half of the aperture is completely dark or bright. The limits of movement of the rear mirror at which the fringe no longer fills the complete field of view are also established. After these limits are established, the mirror is again moved to the position of zero path difference and moved to the closest position at which the field is at its maximum brightness. If now a short focal-length eyepiece lens is used to focus on the 5-mm square aperture, the razor blade in the right arm of the interferometer will cast a shadow in this aperture and diffraction bands will be seen if this edge is nearly parallel to the horizontal slit of the original illuminated cross. This razor blade should then be rotated until these fringes are sharpest and raised or lowered until the fringes appear in the center of the 5-mm square aperture. This adjustment makes it certain that the plane waves from the upper and lower halves of the aperture, which arise from the horizontal line of the illuminated cross, will not on their path to the focal plane cross each other. This is necessary if our original conditions of Fig. 1 and Eq. (20) are to be met. The plane waves arising from the vertical line of the illuminated cross will cross each other, however, but this in no way disturbs the form of the vertical diffraction pattern. A slight rotation of the diffracting aperture about a horizontal axis will make these diffraction bands parallel to the upper and lower edges of the aperture. The razor blade in the rear interferometer arm is then inserted and by viewing the diffraction bands produced by this straight edge, a similar adjustment of parallelism and height in field of the 5-mm square aperture can also be made. When these two sets of diffraction bands are adjusted so that the positions corresponding to the geometrical shadow of each straight edge are overlapping and parallel to each other, the adjustments are complete. The long focal-length lens is then replaced in its proper place and the image of the cross may either be brought to a

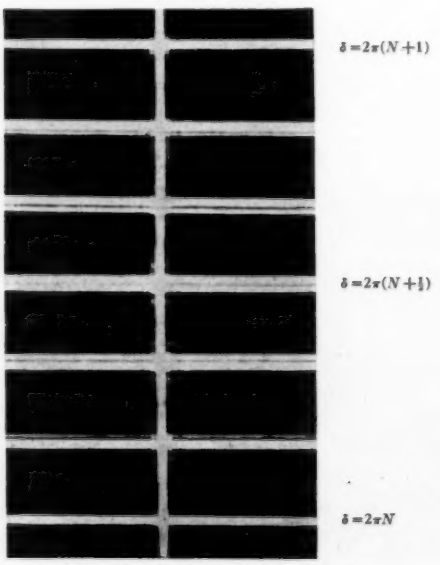

FIG. 15.

focus on a film or viewed with a short focal-length eyepiece.

Assuming that the polarizer is in place between the source and the illuminating cross and is oriented in the desired direction, the previous setting of the instrument just described should give a diffraction pattern corresponding to Eq. (20) when $\delta=0°$ (or more generally $2\pi N$). The vertical and horizontal diffraction patterns should then be of the same form as those shown in Figs. 4 and 5. A series of pictures were then taken during which the rear interferometer mirror was moved by some small, previously determined amount before each exposure. When the complete cycle was covered (upper half of the field of view bright again when the rear razor blade is removed from the field of view), the resulting pictures were enlarged about four times to obtain the series of pictures shown in Figs. 15 and 16. The series of pictures in Fig. 15 are somewhat underexposed to emphasize the structure of the central bands while those in Fig. 16 are a little over-exposed to show the structure of the smaller and weaker secondary diffraction bands. If the patterns corresponding to $\delta=2\pi N$ and $\delta=2\pi\ (N+1)$ in each of the two figures are observed, it can be seen that the vertical and horizontal bands in each figure have the same shape as was predicted theoretically in Figs. 4 and 5. Going upward in Fig. 15, a satellite appears on the upper side of the horizontal line grows until at $\delta=2\pi(N+\frac{1}{2})$ both lines are of the same intensity which again was predicted theoretically by Fig. 10. Going upward in the series still further, we see that the satellite now becomes the stronger line and at $\delta=2\pi(N+1)$ the pattern is the same as the original one. The vertical band system throughout the sequence has remained unchanged as was predicted by Eq. (20) and Fig. 4. Going upward similarly in Fig. 16, the secondary band structure in both vertical and horizontal directions starts out having the same form as again predicted in Figs. 4 and 5. At $\delta=2\pi(N+\frac{1}{2})$ alternate bands in the horizontal direction have disappeared while the remaining ones are much enhanced which again was predicted by Fig. 10. At the top picture where $\delta=2\pi(N+1)$, the pattern is the same as that at the bottom. Again throughout the sequence of pictures it can be seen that the vertical band system has remained unchanged. The actual positions of the bands on the film or plate for any given value of δ can be predicted from Eqs. (20) and (11) when once the geometry of the system (a, b, F) and the wave-length of the light λ are known.

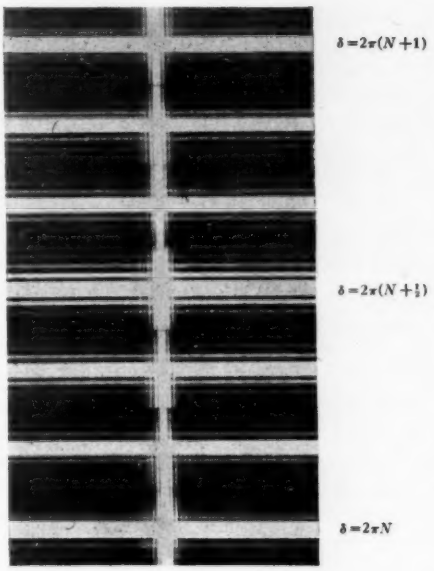

FIG. 16.

Although Figs. 15 and 16 show the forms of the bands for varying values of δ, no attempt has been made to show the movement of the fringes which are responsible for these changes. Since the forms of the vertical diffraction bands in all cases remain unchanged, we will be concerned here only with the changes in the horizontal band structure. To attempt to show the fringe movement, horizontal parts of the patterns in Fig. 15 and 16 were cut out and mounted in order side by side so as to show the progressive movement of the bands as the rear interferometer mirror is

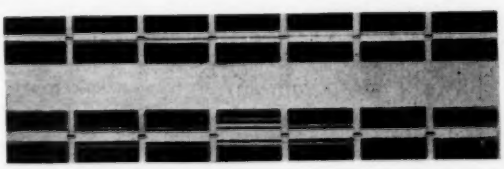

FIG. 17.

moved constantly in the same direction (see Fig. 17). The dark horizontal ink line just visible between adjacent photographs serves as a fixed line of reference. The diffraction patterns were so arranged about this line as to indicate the progressive movement of the bands as the photographs are viewed in sequence going from left to right. In the upper series, which shows the movement of the central bands, the fringe movement is progressively upward in going from left to right. A weak satellite first appears on the lower side of the present line increases in intensity and moves upward. The parent line simultaneously decreases in intensity and moves upward. At the central photograph the satellite has the same intensity as the parent line and both are symmetrically arranged about the fixed line of reference. In the photographs to the right of the central one, the satellite while still moving upward in the field now becomes the stronger line and continues to increase in intensity while the parent line still decreasing in intensity also continues to move upward in the field. At the extreme right the satellite occupies the same position and has the same intensity as the original parent line. The changes in the smaller diffraction bands can also be seen from the lower series of photographs in Fig 17. In the photograph at the left there will be seen to be present a series of parallel, equally spaced, finer diffraction bands which have the same spacings as the vertical bands in Fig. 16. As one follows the movement of the fringes (again going from left to right in the photographs) it will be seen that all fringes move upward in the field. As has been previously mentioned, however, an additional feature appears in that one set of alternate bands progressively increases in intensity while moving upward and the second set decreases in intensity while moving upward. In the middle photograph of the series, alternate bands have completely disappeared and the remaining bands are more intense and about twice as wide as the original smaller bands. In the photographs to the right of the central position, these stronger bands while still moving upward begin to weaken and become narrower to make room for alternate weaker bands which simultaneously begin to reappear between each of these bands. In the picture at the extreme right, the appearance of the pattern is the same as that at the extreme left. During the cycle, however, if the changes are carefully observed, each fringe will be found to have moved upward one position in the field. These experimental changes in the forms of the horizontal fringes are in complete agreement with Figs. 5–12. Again as shown theoretically, all these changes can be made to take place in a reverse order (fringes will move downward in the field) by simply reversing the direction of the movement of the rear interferometer mirror.

These changes in the diffraction pattern form are quite spectacular to see particularly for the first time. So many changes are taking place at the same time that it is difficult to decide just what is happening. It is only after considerable observation that one arrives at the conclusion that the sequence of events is that described by the foregoing paragraphs and theoretical curves.

After having seen these interesting changes in the diffraction pattern form one immediately tries other experiments. One of the first which suggests itself is that of rotating one of the planes of polarization in the diffracting aperture. Accordingly, therefore, a half-wave plate was inserted over one of the halves of the aperture so that the electric vectors in both regions of the aperture were vibrating perpendicularly to each other. Being sure that the original adjustments of aperture and razor blades were not disturbed, the diffraction pattern was viewed in the focal plane of the long focal-length lens. If, as before, the two interferometer arms are again made nearly equal, it will be found that the diffraction pattern is of the form shown in Fig. 18. Movement of the rear interferometer arm in no way affects the forms of the vertical or horizontal diffraction bands. The secondary horizontal bands have twice the separation of the corresponding vertical bands. These remarks are in complete accord with Eq. (23) and Figs. 4 and 13.

Another experiment which suggests itself is that of rotating the plane of polarization in one-half of the aperture at an arbitrary angle α with respect to that in the other half. According to Eq. (26), the changes which can be expected with increasing values of α between 0° and 90° is a gradual transition from the types of diffraction bands in Figs. 15 and 16, where δ influences the diffraction pattern, to Fig. 18 where δ has no

influence. The forms of these bands were neither calculated nor observed experimentally.

The last experiment which was tried was that of removing the Polaroid and half-wave plate and viewing the diffraction patterns directly in unpolarized light. As might be expected from the previous discussion, this gave nothing new, for the bands were of identically the same form as those shown in Figs. 15 and 16.

In the previous remarks nothing has been said about the possibility of experimentally checking Eqs. (18) and (22) in which waves of different amplitudes are considered. These equations also could be checked by inserting filters of known transmissions over either of the halves of the diffracting aperture. These filters ought also to be of interferometer quality and the phase retardation introduced by their presence also has to be known. No attempt, however, was made experimentally to do this.

Since in these experiments the form of the resultant diffraction patterns can be altered, it follows directly that the resolving power of a telescope in the horizontal and vertical directions can be changed. If, therefore, a telescope were directed into the interferometer and a distant resolving-power chart viewed through the combined system of interferometer and telescope, the resolving power in the vertical direction will remain constant while that in the horizontal direction can be changed at will by as much as a factor of two by simply rotating the interferometer screw. As far as the author is aware also these experiments are unique in that they make it possible to measure linear distances in terms of diffraction patterns instead of the usual interference methods. Thus if a good plane parallel plate were coated over half its surface with a half wave-length retardation film, it should be possibly by viewing (preferably microphotometering) the diffraction pattern before and after the plate was inserted in front of the diffracting aperture to determine the optical thickness of the film. In making these observations, the dividing line between the coated and uncoated regions should be placed directly over the dividing line of the aperture. If the diffraction pattern before inserting the plate was of the form of the middle photograph in Figs. 15 and 16, after insertion of the plate, the diffraction pattern would change its shape to the upper or lower diffraction pattern form. Linear distances can also be measured in terms of light waves by counting the number of times a given diffraction pattern is repeated when moving the rear interferometer mirror through some unknown distance and remembering that the displacement of the mirror necessary to repeat the pattern once is $\lambda/2$.

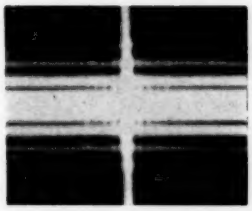

FIG. 18.

The author is indebted to several who helped to put this manuscript in its final form. Dr. G. B. Sabine of our Division read the complete manuscript and made some suggestions which added to its clarity and continuity. The authors wife, Dr. Frances F. Mahan also aided materially in the preparation of the paper. The author is further indebted to Miss Betty Hechmer, and Miss Alice Crosby of the Naval Ordnance Laboratory, who prepared the many diagrams.

Spark Light Source of Short Duration*

J. W. Beams, A. R. Kuhlthau, A. C. Lapsley, J. H. McQueen, L. B. Snoddy, and W. D. Whitehead, Jr.
University of Virginia, Charlottesville, Virginia

(Received July 7, 1947)

A simple spark-gap light source together with an associated pulsing circuit has been utilized to produce light flashes of approximately 10^{-7}-sec. duration suitable for the study of shock waves produced by high speed bullets and for the photography of shock and turbulent disturbances in supersonic air streams.

THE detail revealed in schlieren and shadow photographs of rapidly moving objects depends upon the total time of illumination of the object. The shorter the light flash the better the detail, provided that light energy sufficient to activate the photographic emulsion is available. In the study of turbulence and phenomena in boundary layers very short exposure photographs are certainly of value, and motion pictures at a high repetition rate with very short exposure times are of particular interest. With these advantages in mind a source has been devised which produces light flashes of approximately 10^{-7}-sec. duration and has sufficient intensity for ordinary photographic purposes.

The source consists of a conventional spark gap operating in air at atmospheric pressure, which is energized by a pulsing circuit capable of supplying high current for a short time and then reducing the current to zero, or to a very small value.

It is well known[1] that in the ordinary type of spark discharge the spectral lines due to the gas (principally spark lines when air is used) appear first, followed by the spark and arc lines characteristic of the electrode material. The spectral lines of the gas are emitted with relatively high intensity during the high potential stage of the discharge before the low voltage arc is formed. That is, the first light emitted by the spark originates in the gas surrounding the electrodes. As the discharge progresses, vapor of the electrode material is formed on the electrodes and moves toward the center of the spark gap. This vapor gives rise to the spark and arc lines characteristic of the electrode material. At small gap spacings and minimum field strengths the transition from the high potential to the arc discharge takes probably from 1 to 5×10^{-8} sec. This time is not to be confused with the lag of the spark, but refers to the time taken by the discharge when once started to change from a high potential to a low potential arc type. This time can, of course, be made much longer by increasing the electrode separation, using extremely non-uniform fields, etc.,[2] but for two electrodes with a reasonably uniform field distribution and a gap spacing of one centimeter, it can be made less than 5×10^{-8} sec.

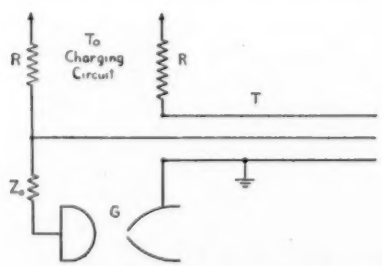

FIG. 1. Schematic diagram of spark-gap light source, G, and associated coaxial transmission line, T, for producing short current pulse. R—charging resistances. Z_0—resistance equal to surge impedance of line T.

In this work an attempt was made to reduce the current through the gap to zero or to a very small value as soon as the arc-type of discharge was formed and consequently limit the length of the current pulse to the high voltage transition period. After the removal of the external field there are, of course, various decay processes which prolong the light emission. However, these have not proved to be serious provided light flashes of 10^{-7}-sec. duration are considered satisfactory.

* This work was supported by the Bureau of Ordnance of the Navy under Contract NOrd-7873.

[1] J. W. Beams, Phys. Rev. **35**, 24, (1930).

[2] L. B. Snoddy, J. R. Dietrick, and J. W. Beams, Phys. Rev. **52**, 739 (1937).

The transition period probably could be decreased and the light emission increased by overvolting the gap. It is known[3] that the field strength in a gap in air at atmospheric pressure can be maintained at about 16 times the minimum value for periods up to 10^{-6} sec. without breakdown provided the gas surrounding the electrodes is free from dust nuclei, ions, etc. If during this period of high field the discharge is initiated by the liberation of an electron at the cathode surface, the light intensity should be greatly increased and the transition period decreased. While this gap would seem to have decided advantages, it was not necessary to use it in this preliminary work.

The circuit used is shown in Fig. 1. It consists of a transmission line, open ended at the output end, and shunted at the input end by a spark gap in series with a resistance equal to the surge impedance of the line. The line is charged initially from a conventional rectifier circuit with the charging current limited by high resistances to a few milliamperes. For random flashes, the line is charged slowly until the spark gap discharges. If it is necessary to have the gap discharge at approximately constant potential the gap electrodes should be made of photosensitive material and irradiated by ultraviolet light.

The operation of the circuit is simple in principle. When the spark gap breaks down a discharge wave travels from the input to the output end, is reflected at the output end, and returns to the input end reducing the potential of the line to zero. This of course implies that the spark-gap resistance has decreased to a value low compared to the characteristic impedance of the line by the time the reflected wave returns to the input end. If this is not the case, there will be reflection at the input end and the discharge will be prolonged. This means that the time length of the line should be at least as long as the transition period of the discharge. To increase the intensity of the light two or more transmission lines can be connected in parallel using one gap to discharge the combination. In this case, each line should be connected separately to the gap electrodes through a resistance equal to its characteristic impedance. It is extremely important to keep the length of the connecting wires at an absolute minimum and to use non-inductive resistors in the discharging circuit. The gap used in this work was the conventional cone type with the discharge viewed in the axial direction through a small hole in one electrode.

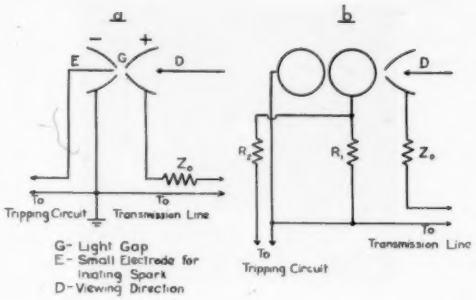

FIG. 2.

If it is necessary to synchronize the light flash with other events, the gap can be tripped by any of the usual devices. Two possible arrangements are shown in Fig. 2. With either method the flash can be started within 10^{-6} to 10^{-7} sec. of any desired time. However, to produce accurate tripping the synchronizing potential pulse must have a steep front and an amplitude sufficient to raise the field strength in the gaps considerably above the minimum sparking value.

With a transmission line of about 10^{-7}-sec. electrical length (i.e., 10^{-7} sec. is the time required for the discharge wave to travel from input end to output end and back to input), the

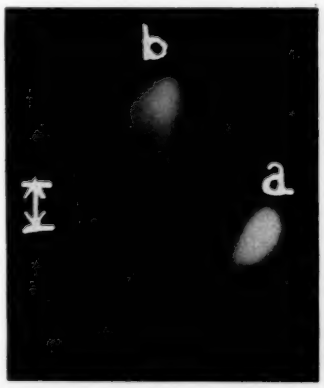

FIG. 3.

[3] J. C. Street and J. W. Beams, Phys. Rev. **38**, 416, (1931).

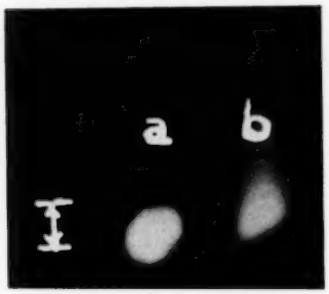

FIG. 4.

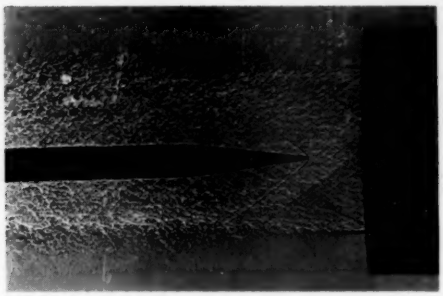

FIG. 5. Metal probe, diameter $\frac{3}{16}$", in supersonic free jet. Mach number approximately 1.8.

duration of the flash was determined by the use of a high speed rotating mirror. The mirror was carried by an air turbine and rotated at 1500 r.p.s. The mirror-to-image distance was 9 feet.

The duration of the flash was determined by allowing the spark image to sweep across a photographic film. Since it is necessary to use a lens in this determination, there might be some question as to whether or not this corresponds to the actual experimental arrangement used in most shadow photography. In order to investigate this point, photographs were taken using quartz and glass lenses. Figure 3 was taken with a quartz-fluorite combination and Fig. 4 with a corrected glass lens. Two pictures are shown on each film, (*a*) with the mirror stationary and (*b*) with the mirror revolving at 1500 r.p.s. The distance marked by the arrow in each case represents 10^{-7} sec. In neither set of photographs does the light have any appreciable effect after 10^{-7} sec.

The electrical constants of the circuit are given below:

$R = 2 \times 10^6$ ohms,
$Z_0 = 52$ ohms,
gap voltage = 20,000 volts,
diameter of hole in light gap = 0.020 in.

The transmission line T consisted of 9 feet of cable An no. RG8/U. Actually two lines were connected in parallel for most of these experiments.

This source has been used in our laboratory for the photography of various phenomena. The detail revealed by using a light source of short duration is well illustrated in Fig. 5 which is a typical photograph of the disturbances in a supersonic free jet. The nozzle used in this experiment was designed for a Mach number of 1.83.

It is clear that the above type of electrical circuit in various modifications can be applied to other than the simple spark-gap light source described above. Also that, if necessary, the length of the light flash produced can be made much longer by employing an artificial instead of a real transmission line.

The Decoloration of Natural Fluorite Crystals*

ROBERT HERMAN AND SHIRLEIGH SILVERMAN
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received July 14, 1947)

The simple heat treatment for decoloring fluorite crystals of various hues is described. The improvement in transmittance is shown by means of spectrophotometer curves.

INTRODUCTION

WHILE the thermoluminescence of fluorite and other minerals is well known,[1] there seems to have been little attention given to the decoloration of fluorite by thermal treatment. This effect has some practical importance to the optical industry in that an appreciable amount of flawless natural fluorite must be rejected because of the presence of absorption bands in the visible and near ultraviolet.[2] Coloration of various hues occurs and in all cases makes the mineral unsuitable for optical instrumentation. Since it is well known that many varieties of fluorite progressively lose their color as they thermoluminesce it seemed worth while to study the absorption spectra of colored single crystals of natural fluorite before and after heat treatment.

EXPERIMENTAL PROCEDURES AND RESULTS

Unless precautions are taken to avoid thermal shock, fluorite crystals decrepitate badly during heat treatment. Accordingly, samples of colored single crystals of fluorite, up to 15–20 cm^3 in volume, were placed in a thick-walled steel container which in turn was placed in an annealing oven. The temperature was raised from 20°C to about 300°C at the rate of approximately 50° per hour. The temperature was maintained at the higher level for 72 hours and then brought back to room temperature over a period of about 24 hours. In this way cracking of the crystals was kept to a minimum and a number of fairly large plates of excellent clarity were readily obtained. It was also found, as expected, that higher temperatures produced decoloration as well as thermoluminescence at a greater rate. However, fracturing was noticeably worse the higher the temperature at which annealing was carried out.

Transmittance curves for several polished sections of fluorite were obtained in the range from 200mμ to 800mμ on Beckman and G. E. spectrophotometers. The transmittance was measured at intervals of 20mμ, except in the region of absorption bands where some of the measurements were taken at intervals of 10 and sometimes 5mμ. Curve *A* in Fig. 1 shows the transmittance of a polished section of pale purple fluorite, 2.77 mm thick. There are three prominent bands appearing at about 580, 395, and 335mμ. Curve *B* in Fig. 1 was obtained with a 1.22-mm polished

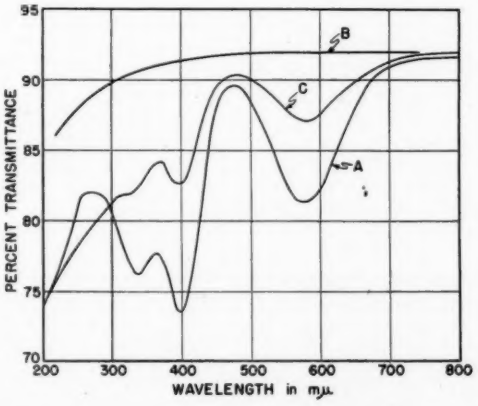

FIG. 1. Transmittance curves for pale purple natural fluorite. Curve *A*—polished plate of pale purple fluorite, 2.77 mm thick. Curve *B*—polished plate of decolored pale purple fluorite, 1.22 mm thick. Curve *C*—polished plate as in curve *B*, after irradiation with x-rays.

* A portion of the work described in this paper has been supported by the Bureau of Ordnance, U. S. Navy, under Contract NOrd-7386.

[1] An excellent bibliography on the subject of thermoluminescence is to be found in M. A. Northup and O. I. Lee, "Experiments on thermoluminescence of some common and unusual minerals," J. Opt. Soc. Am. **30**, 206 (1940).

[2] Although at the present time there exists a technique for producing synthetic CaF_2, the heat treatment discussed in this note is simple as well as very inexpensive and may render available for optical purposes a reasonable fraction of what is now inferior fluorite.

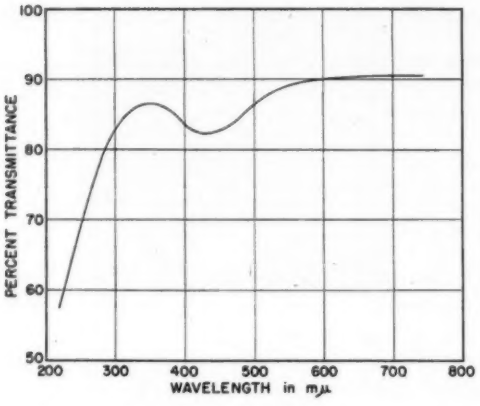

FIG. 2. Transmittance curve for a 1.22-mm thick polished plate of amber-colored fluorite.

section of a decolored crystal that had originally been of the pale purple variety. The improvement in transmittance is striking and takes place over most of the spectral range examined. The decolored crystal was irradiated with 150 kv x-rays. After a period of 6 hours the crystal exhibited the original purple hue. This can be seen by the inspection of curve *C* in Fig. 1 where the same absorption bands have developed as were present originally in the natural purple colored variety. In Fig. 2 is shown the transmittance curve for a 1.22-mm thick polished plate of fluorite which is amber colored because of the broad absorption band at about 430mμ. It should be pointed out that this colored variety as well as others can be decolored quite successfully. Some of the reduction in transmittance for the samples shown in Figs. 1 and 2 may be due to scattering by striae. The above results justify the comments of Mr. A. H. Bennett of the American Optical Company, Scientific Instrument Division, to the effect that the decolored fluorites more than met the color requirements for microscope lens elements.

ACKNOWLEDGMENTS

The authors wish to express their thanks to Mr. A. H. Bennett of the American Optical Company, Scientific Instrument Division, for the gift of several samples of fluorite, to Dr. Archie Mahan of the Naval Ordnance Laboratory for his interest in this problem, and to Dr. K. S. Gibson of the National Bureau of Standards for the use of spectrophotometeric equipment. Thanks are due Mr. H. J. Keegan and Miss M. A. Balknap also of the National Bureau of Standards for their cooperation in obtaining the transmittance curves.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 10 OCTOBER, 1947

Coating Thickness Measurement by Interferometry*

M. F. BECHTOLD
Du Pont Experimental Station, Wilmington, Delaware

(Received May 15, 1947)

Thickness measurements of thin transparent coatings on thick transparent bases by reflection interferometry are made possible through use of an immersion medium of the same refractive index as the base to intensify interference. A convenient interferometer comprising three essential parts, a "pocket" spectroscope, an incandescent lamp, and a cylindrical immersion tank, is described. Its use for rapid, precise measurements of thickness of single and double layer coatings and films in the range of 0.2 to 50 microns is illustrated with spectra.

I. INTRODUCTION

THE need for rapid, accurate measurements of the thickness of thin transparent laminae has long been apparent to those concerned with the physics and chemistry of surfaces or the production and evaluation of protective coatings and unsupported films. Thicknesses of 0.2 to 10 microns, which are below the sensitive range of ordinary mechanical calipers, are particularly difficult to measure. The tedious methods of stripping or sectioning followed by microscope examination as well as the limited methods of x-ray,[1] ultraviolet, and visible light absorption have been used for this purpose. Recently, the thickness of thin silvered films on glass has been measured by multiple-beam wedge interference.[2,3] The ordinary reflection interference method[4,5] for the measurement of unsupported films and the immersion technique of Pfund for showing interference due to coatings[6] have been combined in the instrument and procedure described here.

II. THEORY

Analysis of the intensity and phase changes[7] in light reflected from the two interfaces of a transparent lamina of refractive index μ_1, immersed in a medium of refractive index μ_0, shows that interference minima will occur when the optical path difference of the two reflected beams is equal to a whole number of wave-lengths, or

$$2\mu_1 t\cos\theta = n\lambda, \tag{1}$$

where t is the thickness, θ is the angle (to the normal) of refracted light in the lamina, n is an integer, and λ is the wave-length in air.

Thus, if white light is incident on the lamina, a number of dark bands appear in the spectrum of the reflected beam, the number and location of which can be predicted as follows: let ϕ be the angle of the incident (or reflected) beam to the normal; then from Snell's law,

$$\sin\theta = (\mu_0/\mu_1)\sin\phi, \tag{2}$$

and from a trigonometric relationship,

$$\cos\theta = \left[1-\left(\frac{\mu_0}{\mu_1}\sin\phi\right)^2\right]^{\frac{1}{2}}. \tag{3}$$

Hence, from Eq. (1),

$$2\mu_1 t\left[1-\left(\frac{\mu_0}{\mu_1}\sin\phi\right)^2\right]^{\frac{1}{2}} = n\lambda. \tag{4}$$

For the usual case in which two or more interference bands are visible in the spectrum at one value of ϕ, i.e., one viewing angle, let

$$k = 1\Big/\left\{2\mu_1\left[1-\left(\frac{\mu_0}{\mu_1}\sin\phi\right)^2\right]^{\frac{1}{2}}\right\}; \tag{5}$$

then, from Eq. (4),

$$t/k\lambda_i = n_i \tag{6}$$

* Contribution No. 220 from the Chemical Department, Experimental Station, E. I. du Pont de Nemours and Company, Inc.

[1] H. Friedman and L. S. Birks, Rev. Sci. Inst. **17**, 99 (1946).

[2] Gunn and Scott, Nature **158**, 621 (1946).

[3] Plessner, Nature **158**, 915 (1946).

[4] R. W. Wood, *Physical Optics* (The Macmillan Company, New York, 1934), third edition, p. 192.

[5] H. W. Straub, Alien Property Custodian No. 334, 220 (May 4, 1943).

[6] A. H. Pfund, J. Opt. Soc. Am. **36**, 95 (1946).

[7] Jenkins and White, *Fundamentals of Physical Optics* (McGraw-Hill Book Company, Inc., New York, 1937), first edition, p. 83.

for any selected band, i. Also let

$$t/k\lambda_s = n_s \tag{7}$$

for any other band, s, where $\lambda_i > \lambda_s$. Now by subtraction,

$$(n_s - n_i) = t/k[(1/\lambda_s) - (1/\lambda_i)], \tag{8}$$

or

$$t = (n_s - n_i)k[(\lambda_s\lambda_i)/(\lambda_i - \lambda_s)]. \tag{9}$$

The quantity $(n_s - n_i)$ is obviously equal to the total number of bands from λ_s to λ_i, inclusive, minus one. Thus, when only two adjacent bands are used for calculation,

$$t = k[(\lambda_s\lambda_i)/(\lambda_i - \lambda_s)]. \tag{10}$$

In the case of a coating, the refractive index change at the coating/base interface is fixed. Consequently, in order to obtain an equivalent change at the immersion medium/coating interface, an immersion fluid of refractive index μ_0, equal to that of the base, is chosen. This has the effect of equalizing the intensity of the total beams reflected from the two interfaces so that the interference is complete. The location of interference bands viewed at $\phi = 45°$ due to a coating of $\mu_1 = 1.46$ on a base of $\mu_0 = 1.50$, immersed in a liquid of $\mu_0 = 1.50$, is shown graphically in Fig. 1. The value of k is 0.50; the effect of dispersion is neglected. The family of lines shown were obtained by plotting t *vs.* λ from Eq. (6), letting n vary from 1 to 15. Their intersections with the dotted horizontal lines show, for example, that when $t = 3$ microns, six interference bands occur within the easily observable part of the visible spectrum. (It is difficult to observe bands outside the range of 0.425 to 0.695 micron for λ in the apparatus described later.) The number of bands observed does not decrease with thickness in a regular manner. For example, at $t = 1.03$ micron, only one interference band should be easily observable, while a thinner coating (0.93 micron) should show two bands. A coating of $t = 0.36$ micron should give no bands at all, while a thinner coating (0.27 micron) should give one band in the middle of the spectrum. However, for the 1.03- and the 0.36-micron coatings, the use of a slightly different value of ϕ would be adequate to bring two bands into the observable spectrum to provide data for accurate calculation of t.

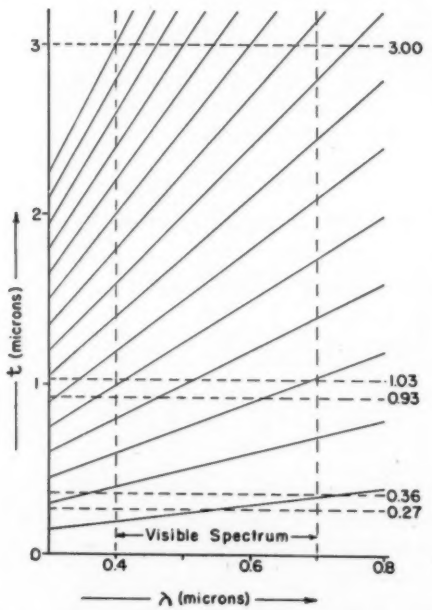

FIG. 1. Location of interference bands in the visible spectrum of light reflected from a coating of refractive index 1.46 (μ_1) on a base of refractive index 1.50 (μ_0), viewed at an angle of 45° (ϕ). Intersection of the family of sloping lines ($t = nk\lambda$) with horizontal lines shows value of λ at reflection minima.

Rigorous treatment of the problem of simultaneously measuring the thicknesses (t_1, t_2) of two adjoining laminae would require consideration of the fact that the beam transmitted by the first lamina encountered contains slight intensity maxima and has suffered some dispersion. Furthermore, the angle of incidence for the second lamina (ϕ_2) is equal to angle of refraction in the first (θ_1), and the condition for minima for either layer may involve $(n+\frac{1}{2})\lambda$ instead of $n\lambda$, dependent on the phase changes at its two interfaces. No attempt has been made to incorporate these complicating factors into general equations; however, calculations of t_2 have been made in this paper using a value of ϕ_2 for the second layer calculated from the values of μ_1, μ_2, and ϕ_1. Examination of Eq. (10) shows that accuracy in the value of t depends chiefly on the number and location of the bands, while it is not sensitive to small inaccuracies in estimates of refractive index and of ϕ at less than about 50°. Consequently, when μ_0, μ_1, and μ_2 are nearly equal, and the same ϕ may be used for both layers,

good approximations of t for both layers can be made with the one value of k. The two sets of bands predicted for the case of one thin layer and one thick layer should be easily separable by differences in their sharpness and number per $\Delta\lambda$. Since the intensity of the reflected ray and the completeness of interference is different for each layer when μ_0 is not equal to either μ_1 or μ_2, sets of bands due to coatings nearly alike in thickness should be separable by intensity differences alone. When the refractive index situation is favorable, a third set of bands should be visible, which arise from interference in reflections from the top and bottom interfaces of the double coating.

III. APPARATUS AND PROCEDURE

An apparatus for observing the interference bands in light reflected at various angles is shown in Fig. 2, and a simplified diagram of the optical elements is shown in Fig. 3. The three filaments of a 200-watt 110-volt photo-projection bulb are placed in line with a slit (6 mm wide) and the center of an immersion tank at a distance of 21 cm. The light is reflected from the surface of a coating or film held at any angle by the revolving clamp on top of the immersion tank (7-cm O.D.), and the spectrum of the reflected light is analyzed with the pocket spectroscope, which has been rotated to the appropriate angle. To provide for this rotation, the spectroscope is mounted on a wheel of polymethyl metacrylate sheeting that rotates about the immersion tank as an axle. The tank does not turn, since its bottom is slotted and keyed to a brass base plate. A fixed wheel of polymethyl methacrylate bearing calibrations in terms of the angle ϕ permits easy measurement of the position of the spectroscope on the wheel above it. Thus, the apparatus provides for independent rotation of the sample and the spectroscope. The clamp positions the sample so that the incident beam, which enters the cylinder along a radius, is reflected at the axis of the cylinder along another radius to the spectroscope without refraction such as would be encountered with a tank of rectangular cross section.

The spectroscope used for all results given here was manufactured by Zeiss (No. 71112). It is especially convenient since it contains a readily

FIG. 2. Interferometer.

calibrated wave-length scale; illumination for this scale is furnished by oblique light from the source reflected from the exterior of the immersion tank down the side arm of the spectroscope for angles less than 45°. A small mirror (not shown) is used for higher angles. The spectroscope is positioned with the entrance window about 1 cm from the immersion tank and is used substantially fully telescoped. A further advantage of the Zeiss spectroscope for this use is that a reference spectrum, such as a mercury arc, can be introduced through a side prism and placed in juxtaposition with the experimental spectrum for comparison purposes. A student-

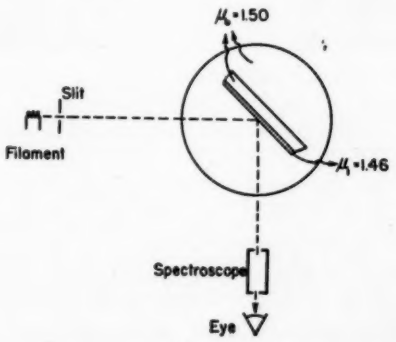

FIG. 3. Optical elements of interferometer.

type diffraction grating spectroscope is also satisfactory for thickness measurements.

Non-colored photographs of the spectra were made by placing a shielded photographic plate at distance of 6.5 cm back of the spectroscope eyepiece with the telescope setting at about 13, the slit opening in the range of 0.4 to 2.5, and exposure time of 15–40 seconds. A "Daylite" filter was used (only for pictures, except Fig. 4C) between the source and the immersion tank in order to lower the intensity of light in the yellow-red end of the spectrum; conditions were such that the intensity minima of the filter did not interfere with the interference spectra. Considerable contrast visible to the unaided eye was lost in the photographic process; in addition, the wave-length scale could not be brought into focus and was blocked off during exposure. Consequently, it was much easier to observe the spectrum with the unaided eye, and to make calculations therefrom, than it was to read the data from photographs. The wave-length data accompanying Fig. 4 were obtained by reading the wave-length scale with the spectroscope telescoped directly before and after taking of the pictures.

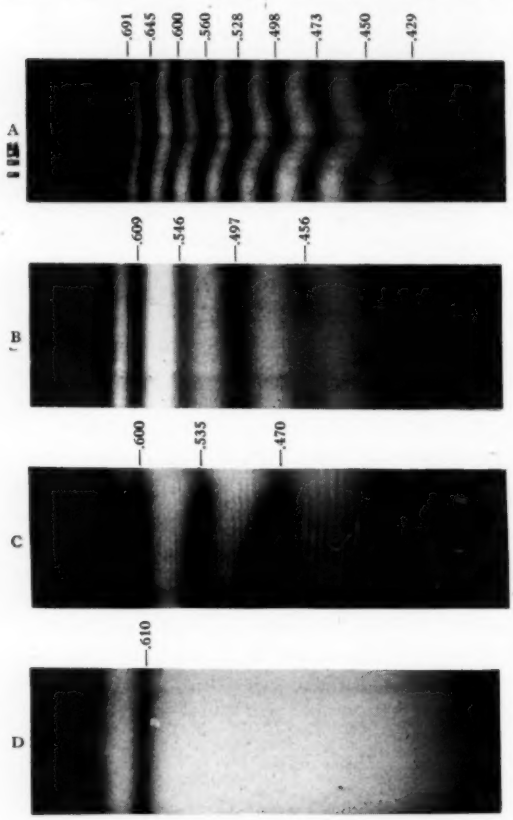

FIG. 4. Photographs of reflection spectra of various laminae.

A. Polyallyl methacrylate (μ_1=about 1.51), on a thick polymethyl methacrylate base (μ_0=1.50) at ϕ_1=47.5°. (k_1=0.487; t_1=4.41 microns.)

B. Two layers of a synthetic resin (μ_1=about 1.46, μ_2=about 1.48), on a thick polymethyl methacrylate base, (μ_0=1.50) at ϕ_1=47.5°. (k_1=0.525; t_1=1.56 micron; k_2 =0.452, t_2=1.25 micron.)

C. One layer of a synthetic resin (μ_1=about 1.48), on Cellophane (μ_2=about 1.52), in a liquid (μ_0=1.52) at ϕ_1=52.5°. (k_1=0.582; t_1=2.53 microns; k_2=0.518, t_2=23.6 microns.)

D. Polystyrene (μ_1=1.59), on a thick polymethyl methacrylate base (μ_0=1.50) at ϕ=53.5°. (k=0.478; t_1 =0.29 micron.)

When an estimate of refractive index of a lamina from its composition is not sufficiently accurate, and the lamina is available in unsupported form, the apparatus of Fig. 2 can be used for the measurement by choice of an immersion liquid (such as a mixture of Nujol and Aroclor) so that minimum light is reflected from the surface down the properly aligned spectroscope. (A refined procedure for this purpose has been described by Billmeyer.[8]) Independent measurement of the immersion liquid establishes the surface refractive index of the lamina. An apparatus for measuring the Brewster angle[9] also provides data for calculation of the surface refractive index.

IV. RESULTS

Calibration

A calibration experiment was performed by mounting a section of a glass bubble (μ_1=1.52) on a frame and observing the location of interference bands when examined at two angles with air (μ_0=1), then water (μ_0=1.33) in the immersion tank. Thus, the glass film is found to average 1.35 micron thick, independent of the immersion medium or the angle of viewing. This experiment and others show that at this thickness the average deviation of a single determination from the mean is about 2 percent.

[8] F. W. Billmeyer, Jr., Phys. Rev. **71**, 489 (1947).
[9] A. H. Pfund, J. Opt. Soc. Am. **31**, 679 (1941).

Non-Uniform Coatings

Ordinary techniques of applying coatings, such as dipping, spraying, or puddling yield obviously non-uniform coatings, since the interference colors observed upon immersion are not the same in all areas. When the coatings are too thick to show interference colors, there has previously been no convenient method of determining the presence of non-uniformity of thickness. A spectrum indicating the presence of a mild non-uniformity in a coating of polyallyl methacrylate on polymethyl methacrylate within the area viewed with the spectroscope is shown in Fig. 4A. In other areas the changes in thickness varied from none at all to complete absence of the coating. The thickness in the center of the spot examined is calculated to be $8\times0.487\times(0.691\times0.429)/0.262=4.41$ microns.

Two-Layer Coatings

An apparently uniform two-layer coating was prepared on a polymethyl methacrylate panel by constant-speed withdrawal, then drying, after dipping in solutions of synthetic resin of slightly different composition. The spectrum of this coating (Fig. 4B) shows two sets of two interference bands of distinctly different intensity. It was established by separate experiments on similar coatings in single-layer form that the lighter set of bands is caused by the layer next to polymethyl methacrylate, and the darker set by the top layer. In addition, the sum of thicknesses measured in single-layer form was equal to the total thickness in two-layer form. Calculation from Fig. 4B gives $t_1=0.525(0.609\times0.497)/0.102=1.56$ micron for the top layer and $t_2=0.452(0.546\times0.456)/0.090=1.25$ micron for the lower. In coatings of this kind, cases of partial and complete coincidence of the two sets of bands have been observed. The failure to observe a third set of bands corresponding to the over-all coating thickness may be due to too close approach of the refractive index of the bottom layer to that of the base, resulting in poor equalization of intensity of the two beams involved.

Another example of measurement of thickness of two laminae simultaneously is seen in Fig. 4C, which is the spectrum of a commercial coated Cellophane film. In this case the lines due to the coating are obviously the three broad ones, while the Cellophane film itself (or the Cellophane plus two thicknesses of coating) is responsible for the fine lines. The sample of coated film, which was wet with oil and held clamped between two microscope slides, showed very non-uniform coating thickness; considerable scanning was done in order to find an area giving bands sufficiently straight and continuous to measure. Calculation of coating thickness yields $t_1=2.53$ microns. Although conditions were nearly ideal to prevent interference due to the Cellophane base, since the immersion fluid has nearly the same refractive index as ordinarily ascribed to Cellophane sheeting, there is some doubt as to whether the fine lines observed are due to base or base plus coatings because the Cellophane itself may reflect because of its double refraction. It is thought that in this case, the fine bands are from the base itself, although this point was not proven experimentally. Twenty-two lines were counted in the range of $\lambda=0.65$ to 0.50 micron, equivalent to a thickness of 23.6

In air.

$\phi=45°$; $k=0.372$ λ(microns)	diff.	t(microns)
0.440		
	0.060	1.36
0.500		
	0.080	1.35
0.580		
$\phi=60°$; $k=0.400$		
0.470		
	0.075	1.37
0.545		
	0.105	1.35
0.650		

In water.

$\phi=45°$; $k=0.419$ λ(microns)	diff.	t(microns)
0.440		
	0.075	1.27
0.515		
	0.100	1.33
0.615		
$\phi=60°$; $k=0.504$		
0.435		
	0.080	1.41
0.515		
	0.125	1.33
0.640		

microns. By making use of high refractive index immersion fluids and values of ϕ at 60–75°, all of which tend to make for fewer bands per $\Delta\lambda$, laminae at least 50 microns thick can be measured.

Thin Coatings

For every set of μ_0, μ_1 values, there is a minimum thickness of the lamina which yields a pair of interference bands at a given angle of viewing. Since within the range of 0.7 to 0.4 micron there is no set of numbers bearing the ratio of 2/1, first- and second-order interference bands will never be visible as a pair at one value of ϕ within this range. Consequently, bands for which $n=2$ and 3 constitute the first observable pair. To illustrate use of this analysis, the thinnest coating of polystyrene on polymethyl methacrylate that will show two interference bands at the lowest convenient angle ($\phi=30°$) is 0.43 micron, which gives bands at $\lambda_3=0.4$ micron and $\lambda_2=0.6$ micron. At $\phi=53.5°$, this minimum thickness is 0.57 micron. The order of the single band from a coating of this kind, whose spectrum is shown in Fig. 4D, is in doubt, since if it is of second order, the third-order band may or may not be visible at 0.406 micron. However, rotation to $\phi=48.5°$ caused the single band to shift to 0.665 micron; no other band appeared. If it is of second order, the third-order band should have been easily observed at 0.444 micron at this angle. Consequently, the conclusion is drawn that the band is of first order, and the thickness of the coating is 0.29 micron. In another area on the same coating, the single band was located at 0.665 micron at $\phi=30°$, 0.565 at 45°, and 0.500 at 53.5°, which data yield thickness of 0.24 micron at each position. By observation at several angles and analysis of the data in this fashion, it is possible to ascertain the order and then the thickness of coatings showing single bands down to the minimum thickness showing one band; at $\phi=30°$ this thickness is 0.143 micron when $\mu_1=1.59$, $\mu_0=1.50$.

Acknowledgment

The author is indebted to Professor A. H. Pfund of the Johns Hopkins University for the essential suggestion of the use of an immersion fluid, and to Dr. F. W. Billmeyer, Jr., of the du Pont Plastics Department, who showed that μ_1 of Eq. (1) must be the refractive index of the lamina with respect to air, regardless of the immersion medium, and made the calibration experiment to prove this point. Thanks are extended to Mr. Alexander Marshall, who assisted in the preparation of the photographs.